

# Molecular Phylogeny of the New World Monkeys (Platyrrhini, Primates) Based on Two Unlinked Nuclear Genes: IRBP Intron 1 and $\epsilon$ -Globin Sequences

H. SCHNEIDER, I. SAMPAIO, M.L. HARADA, C.M.L. BARROSO, M.P.C. SCHNEIDER, J. CZELUSNIAK, AND M. GOODMAN  
*Universidade Federal do Para, Centro de Ciencias Biologicas, Departamento de Genetica, Belem, Para, Brazil (H.S., I.S., M.L.H., C.M.L.B., M.P.C.S.); and Wayne State University, School of Medicine, Department of Anatomy and Cell Biology, Detroit, Michigan (J.C., M.G.)*

**KEY WORDS**      Ceboids, Maximum parsimony, Evolutionary models

**ABSTRACT**      Nuclear sequences of the 1.8 kilobase (kb) long intron 1 of the interstitial retinol-binding protein gene (IRBP), previously determined for 11 of the 16 extant genera of New World monkeys (superfamily Ceboidea, infraorder Platyrrhini), have now been determined for the remaining 5 genera. The maximum parsimony trees found, first with IRBP sequences alone and then with tandemly combined IRBP and  $\epsilon$ -globin gene sequences from the same species, supported a provisional cladistic classification with the following clusters. Subtribes Callitrichina (*Callithrix*, *Cebuella*), Callimiconina (*Callimico*), Leontopithecina (*Leontopithecus*) and Saguina (*Saguinus*) constitute subfamily Callitrichinae, and subfamilies Callitrichinae, Aotinae (*Aotus*), and Cebinae (*Cebus*, *Saimiri*) constitute family Cebidae. Subtribes Chiropotina (*Chiropotes*, *Cacajao*) and Pitheciina (*Pithecia*) constitute tribe Pitheciini; and tribes Pitheciini and Callicebini (*Callicebus*) constitute subfamily Pitheciinae. Subtribes Brachytelina (*Brachyteles*, *Lagothrix*) and Atelina (*Ateles*) constitute tribe Atelini, and tribes Atelini and Alouattini (*Alouatta*) constitute subfamily Atelinae. The parsimony results were equivocal as to whether Pitheciinae should be grouped with Atelinae in family Atelidae or have its own family Pitheciidae. The cladistic groupings of extant ceboids were also examined by different stochastic evolutionary models that employed the same stochastic process of nucleotide substitutions but alternative putative phylogenetic trees on which the nucleotide substitutions occurred. Each model, i.e., each different tree, predicted a different multinomial distribution of nucleotide character patterns for the contemporary sequences. The predicted distributions that were closest to the actual observed distributions identified the best fitting trees. The cladistic relationships depicted in these best fitting trees agreed in almost all cases with those depicted in the maximum parsimony trees.      © 1996 Wiley-Liss, Inc.

Traditional classifications employed for neotropical primates (Infraorder Platyrrhini, Superfamily Ceboidea) usually divide them into two families: Cebidae and Callitrichidae (Simpson, 1945; Simons, 1972; Martin, 1990). Cebidae contains all the ceboids except the small-bodied clawed marmosets and tamarins, which are placed in Callitrichidae. *Callimico* (Goeldi's monkey)

is placed in Cebidae because, despite its clawed fingers and small body, Goeldi's monkey has other morphological features that appear more like those of larger-bodied mon-

Received April 21, 1995, accepted December 20, 1995.

Address reprint requests to Dr. Morris Goodman, Wayne State University, School of Medicine, Department of Anatomy and Cell Biology, MRB, 550 E. Canfield Ave., Detroit, Michigan, 48201.

keys than those of marmosets and tamarins. A modified version of this traditional classification has three families: Cebidae (with *Callimico* removed from it), Callimiconidae (restricted to *Callimico*), and Callitrichidae (Hershkovitz, 1977). As these classifications do not attempt to portray strictly monophyletic groups, they do not provide a basis for depicting the phylogenetic relationships which exist among the genera of neotropical primates.

More recently, on the basis of cladistic analyses of morphological characters, Rosenberger (1981), Ford (1986), Rosenberger et al. (1990), and Kay (1990) have agreed that *Callimico* is closer phylogenetically to marmosets (*Callithrix*, *Cebuella*) and tamarins (*Saguinus*, *Leontopithecus*) than to any remaining ceboids. A classification that reflects these findings places *Callimico* together with the four genera of marmosets and tamarins in the subfamily Callitrichinae (Rosenberger, 1981; Rosenberger et al., 1990). In this classification the traditional cebid-callitrichid dichotomy is radically altered: Cebinae, consisting of *Cebus* (capuchin monkeys) and *Saimiri* (squirrel monkeys), is the sister of Callitrichinae and these two subfamilies constitute the family Cebidae. The remaining monkeys, which along with the cebines were members of Cebidae in the traditional classification, group as the family Atelidae at considerable distance from the cebines in this radically altered classification (Rosenberger, 1981; Rosenberger et al., 1990). However the cladograms of Ford (1986) and Kay (1990), if translated into taxonomic classifications, would yield arrangements for the ceboid families that differ from each other and from Rosenberger's.

There is general agreement that the extant ceboids divide up into 16 genera, although some authors (Moynihan, 1976; Rosenberger, 1981, 1984) have suggested that the pygmy marmoset *Cebuella* could be included in the marmoset genus *Callithrix*. There is also agreement that the various ceboid genera form seven monophyletic groups or clades that separated from one another by early to middle Miocene times, i.e., between 23 and 16 million years ago (MYA; Ford, 1986; Fleagle, 1988; Rosenberger et

al., 1990; Kay, 1990). The cladistic classification of Rosenberger et al. (1990) designates for each of the seven clades the lowest ranking taxon that contains all extant members of the clade. One clade, the Callitrichinae, groups together the marmosets, tamarins, and Goeldi's monkeys. Another, the Atelinae, groups together the largest-bodied, prehensile tailed monkeys, consisting of *Ateles* (spider monkeys), *Lagothrix* (woolly monkeys), *Brachyteles* (woolly spider monkeys), and *Alouatta* (howler monkeys). A third clade, the Pitheciini, groups together the specialized seed predators consisting of *Pithecia* (saki monkeys), *Chiropotes* (bearded saki monkeys), and *Cacajao* (uakari monkeys). The remaining four clades consist of single genera: *Cebus*, *Saimiri*, *Aotus* (night monkeys), and *Callicebus* (titi monkeys). In addition, the morphological studies agree not only on the very close relationship between *Callithrix* and *Cebuella* within the callitrichine clade but also on the sister-grouping of *Chiropotes* and *Cacajao* within the pitheciin clade. However, the morphological studies disagree on which two of the four ateline genera are most closely related.

The morphological studies also disagree on the order of cladistic relationship among the seven clades. For instance, Rosenberger et al. (1990) and Ford (1986) place the pitheciins close to atelines (Fig. 1A,B) but Kay (1990) suggests that *Aotus*, *Saimiri*, callitrichines, and atelines are all closer to one another than to the pitheciins (Fig. 1C). Rosenberger et al. (1990) place cebines (*Cebus* and *Saimiri*) with the callitrichines, as already noted, and place an *Aotus-Callicebus* clade as the sister of pitheciins within subfamily Pitheciinae (Fig. 1A). Ford (1980) restricts her subfamily Pitheciinae to the pitheciins (i.e., *Pithecia*, *Chiropotes*, and *Cacajao*) and proposes two different schemes: one has a *Cebus-Saimiri* clade as the sister of all other extant ceboids, and the second has only *Cebus* as the sister of all other extant ceboids with *Saimiri* as the sister of an *Aotus-Callicebus* clade (Fig. 1B). Both of Ford's schemes have Callitrichinae as the sister of her pitheciine-ateline clade. Kay (1990) separates at the base of the tree first *Callicebus* and then *Cebus* from all re-

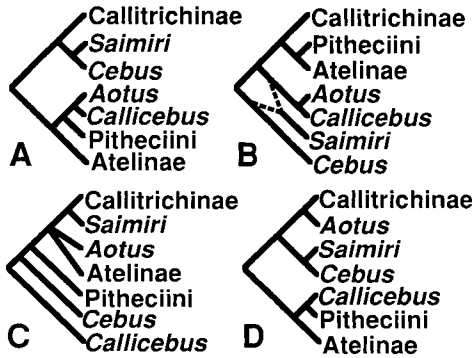


Fig. 1. Alternative views on cladistic relationships among the seven clades that contain the 16 extant ceboid genera. **A:** From the morphological study of Rosenberger et al. (1990). **B:** From the morphological study of Ford (1986). **C:** From the morphological study of Kay (1990). **D:** From analysis of  $\epsilon$ -globin gene sequences (Harada et al., 1995). After Rosenberger et al. (1990), Callitrichinae contains *Callithrix*, *Cebuella*, *Leontopithecus*, *Saguinus*, and *Callimico*; Pitheciini contains *Cacajao*, *Chiropotes*, and *Pithecia*; and Atelinae contains *Lagothrix*, *Brachyteles*, *Ateles*, and *Alouatta*. It should be noted that Ford and Kay use names at higher taxonomic ranks to designate some of these seven clades (e.g., Callitrichidae instead of Callitrichinae).

maining ceboids, has *Saimiri* as sister of the callitrichine clade, and has a trichotomous separation of *Aotus*, the ateline clade, and the *Saimiri*-callitrichine clade (Fig. 1C).

We have confirmed with DNA sequence evidence from two nuclear genes that the 16 extant ceboid genera form the seven clades found in morphological studies. The datasets providing this DNA evidence consist of aligned  $\epsilon$ -globin gene sequences (Schneider et al., 1993; Porter et al., 1995; Harada et al., 1995) and aligned intron 1 sequences from the interstitial retinol-binding protein gene (IRBP; Harada et al., 1995; this study). In humans the  $\epsilon$  gene is part of the  $\beta$ -globin gene cluster, located on chromosome 11 (Bunn and Forget, 1986); IRBP is a single copy gene with four exons and three introns and is located on chromosome 10 (Fong et al., 1990). Presumably these two genes are unlinked in ceboids as in humans, but their chromosomal locations in ceboids have not yet been determined. Our most recent  $\epsilon$  results (Harada et al., 1995) on cladistic relationships within Ceboidea are based on a

dataset of 31 aligned  $\epsilon$ -globin gene sequences from 23 New World monkeys (representing all 16 extant genera) and seven catarrhines (one Old World monkey and six hominoids) and tarsier as the outgroup. This  $\epsilon$  dataset includes sequences from additional *Saimiri* and *Cebus* species over the one *Saimiri* and one *Cebus* sequence previously (Schneider et al., 1993; Porter et al., 1995) representing these genera.

The cladistic results from  $\epsilon$ -globin gene sequences (Harada et al., 1995) for the seven ceboid clades (Fig. 1D) are more similar to the morphological results of Rosenberger et al. (Fig. 1A) than to those of Ford (Fig. 1B) and Kay (Fig. 1C). Fig. 1D differs from Fig. 1A by joining *Aotus* to the callitrichine clade rather than to *Callicebus*, but otherwise has the same branching pattern. These most recent  $\epsilon$  results (Harada et al., 1995), apparently due to their additional *Cebus* and *Saimiri* sequences, joined a *Cebus*-*Saimiri* clade to an *Aotus*-Callitrichinae clade and thus provide a more resolved picture of cladistic relationships within Ceboidea than the previous  $\epsilon$  results (Schneider et al., 1993; Porter et al., 1995), which had depicted an unresolved tetrachotomy of *Aotus*, Callitrichinae, *Cebus*, and *Saimiri*.

Harada et al., (1995) interpreted the maximum parsimony tree for all 31  $\epsilon$ -globin gene sequences as providing strong support (in terms of high bootstrap and parsimony strength of grouping values) for the following monophyletic taxa: Callitrichina (*Callithrix*, *Cebuella*), *Callimico*, *Leontopithecus* and *Saguinus* constituting the subfamily Callitrichinae; Callitrichinae, *Aotus*, *Cebus*, and *Saimiri* constituting the family Cebidae; Chiropotina (*Chiropotes*, *Cacajao*) and *Pithecia* constituting the tribe Pitheciini; Pitheciini and *Callicebus* constituting the subfamily Pitheciinae; Atelini (*Brachyteles*, *Lagothrix*, *Ateles*) and *Alouatta* constituting the subfamily Atelinae; Pitheciinae, Atelinae, and Cebidae constituting the Ceboidea, the only extant superfamily of the infraorder Platyrrhini. In addition, this maximum parsimony  $\epsilon$  tree provides modest support for sister-grouping *Cebus* and *Saimiri* within the cebid clade and weak support for sister-grouping *Brachyteles* and *Lagothrix* within the ateline clade.

Our IRBP dataset originally (Harada et al., 1995) consisted of 16 orthologous intron 1 sequences, representing 15 ceboid species from 11 ceboid genera and the human sequence (Fong et al., 1990) as the only other primate orthologue so far determined and thus the only outgroup to the ceboid sequences. All seven ceboid clades were represented. However the callitrichine clade was represented by just one (*Callithrix jacchus*) of its five genera, and the pitheciine clade by two (*Chiropotes satanus* and *Pithecia irrorata*) of its three genera. The cladistic results from parsimony analysis of these IRBP sequences in most aspects were congruent with the  $\epsilon$  results; the IRBP results only differed from the  $\epsilon$  results by grouping the *Cebus-Saimiri* clade rather than *Aotus* first with the callitrichine clade and by grouping the pitheciine clade with the ceboid clade rather than with the ateline clade (Harada et al., 1995).

In the present study, we have enlarged the IRBP dataset with six more sequences representing two species of *Saguinus* (*S. midas* and *S. bicolor*), *Cebuella pygmaea*, *Callicimico goeldii*, *Leontopithecus rosalia*, and *Cacajao calvus*. Thus, our analysis of IRBP sequences both alone and tandemly combined with  $\epsilon$  sequences now includes representatives of all 16 extant ceboid genera. In addition we use a statistical approach involving evolutionary models (Czelusniak and Goodman, 1995) to further examine the cladistic relationships among ceboid genera. Each evolutionary model consists of an ancestral-descendant branching order upon which a stochastic process of nucleotide substitutions occurs. A goodness of fit statistic represented as loglikelihood ratios, calculated using numerical parameters that minimize these ratios (Powell, 1989a,b), then compares for each evolutionary model its predicted multinomial distribution of nucleotide character patterns in the contemporary sequences to the observed patterns in these sequences. Those ancestral-descendant branching orders in these models giving the closest fit between predicted and observed nucleotide character patterns strongly paralleled the branching orders found in the most parsimonious trees.

## MATERIALS AND METHODS

### Sample description, blood collection and DNA extraction

Blood samples were obtained from the 16 ceboid genera (see Table 1). These animals were captured in different localities of the South American Continent, and nowadays most of them are kept in captivity in primate breeding colonies in Brazil, Argentina, and Peru. The animals were anesthetized with ketalar (10 mg/Kg of body weight) and blood samples were collected through the femoral vein. After centrifugation, white cells were isolated, and DNA extracted following the protocol suggested by Sambrook et al. (1989).

### Primers design and PCR protocols

With primers at 3' and 5' sides of IRBP exons 1 and 2, respectively, a DNA fragment spanning the 1.8–1.9 kb intron 1 of IRBP was amplified by the polymerase chain reaction (PCR), using the following conditions: denaturation 94°C, one minute; annealing: 50°C, one minute; extension: 65°C, three minutes; final extension: 65°C, 10 minutes; 30 cycles. Primers were designed based on the human sequence (Fong et al., 1990). Because this set of primers did not work with callitrichines, new primers were designed based on those ceboid IRBP sequences that had already been determined. Primers based on human sequences (GenBank accession number = JO5253, HUMIRPG) IRB-09 5' CCGCATTCCTGGAATTGTGCCCATGGA-GGT 3' (is the human IRBP nucleotide sequence at positions 3153–3182). IRB-11 5' GCTGCTGGTGGAGCACATCTGGAAGAA-GAT 3' (is the complement of the human IRBP sequence at nucleotide positions 5123–5094). Primers based on atelines and pitheciines sequences: IRB-28 5' ATCCCTGAAATGCCAAGGACCG 3' (correspond to the human IRBP nucleotide sequence at positions 3133–3155). IRB-29 5' GTCCTCAAGCACGTTAGGTTGGA 3' (corresponds to the complement of the human IRBP sequence at nucleotide positions 5031–5009).

### Cloning and sequencing

The amplified DNA fragment was electrophoresed in 1.5% low melting agarose gel,

TABLE 1. List of the ceboid species examined with sequences from  $\epsilon$ -globin (E) and IRBP (I) genes, geographical origins and collecting sites

Species	Common name	Code	Origin <sup>2</sup>	Site <sup>1</sup>	Gene
<i>Alouatta belzebul</i>	Howler monkey	Abe	Rio Tocantins, PA	1	E,I
<i>Aotus azarae</i>	Night monkey	Aaz	Rio Jamari, RO	1	E,I
<i>Aotus nancymai</i>	Night monkey	Ana	Rio Marañon, Peru	3	I
<i>Ateles belzebuth</i>	Spider monkey	Abm	PA	1	E,I
<i>Brachyteles arachnoides</i>	Woolly spider monkey	Bar	Rio de Janeiro, RJ	4	E,I
<i>Cacajao calvus</i>	Uakari monkey	Cca	Loreto, Peru	3	E,I
<i>Callicebus moloch</i>	Titi monkey	Cmo	Rio Tocantins, PA	1	E,I
<i>Callicebus torquatus</i>	Titi monkey	Cto	Unknown	1	E,I
<i>Callimico goeldii</i>	Goeldi's monkey	Cgo	AC	2	E,I
<i>Callithrix jacchus</i>	Common marmoset	Cja	Extremos, RN	1	E,I
<i>Cebuella pygmaea</i>	Pygmy marmoset	Cpy	Unknown	1	E,I
<i>Cebus kaapori</i>	Capuchin monkey	Cka	Rio Gurupi, MA	2	E,I
<i>Cebus nigrovittatus</i>	Capuchin monkey	Cni	Unknown	1	E,I
<i>Chiropotes satanas</i>	Bearded saki	Csa	Rio Tocantins, PA	1	E,I
<i>Lagothrix lagotricha</i>	Wolly monkey	Lla	AM	1	E,I
<i>Leontopithecus rosalia</i>	Lion-tamarin	Lro	Rio de Janeiro, RJ	4	E,I
<i>Pithecia irrorata</i>	Saki	Pir	Rio Jamari, RO	1	E,I
<i>Saguinus bicolor</i>	Tamarin	Sbi	AM	4	I
<i>Saguinus midas</i>	Tamarin	Smi	Rio Tocantins, PA	1	E,I
<i>Saimiri boliviensis</i>	Squirrel monkey	Sbo	Rio Marañon, Peru	3	E,I
<i>Saimiri sciureus</i>	Squirrel monkey	Ssc	Rio Uatuma, AM	1	E,I

<sup>1</sup>1, "Centro Nacional de Primatas", Para, Brazil; 2, Blood collected from animals in the field; 3, "Centro de Reproduccion y conservacion de Primates Nohumanos", Iquitos, Peru; 4, "Centro de Primatologia do Rio de Janeiro", RJ, Brazil.

<sup>2</sup>2 The following abbreviations stand for Brazilian states: PA, Para; RO, Rondonia; AC, Acre; RN, Rio Grande do Norte; MA, Maranhao; AM, Amazonas.

excised and then purified according to the Qiaex gel extraction protocol (Qiagen, Chatsworth, CA). Three different PCR reactions were made for each individual to be sequenced. Each PCR product was purified and the DNA fragment was cloned into the PGEM T-vector from PROMEGA (Madison, WI), following vendor suggestions. Ligation in this vector produced clones at random with distinct orientations, allowing sequencing of both DNA strands. Selected clones were infected with helper phage M13K07 (Invitrogen, San Diego, CA) and the resulting single stranded DNA was purified by PEG-NaCl precipitation followed by phenol-chloroform extraction and ethanol precipitation. The clones were sequenced by the dideoxy chain-termination method (Sanger et al., 1977) using the sequence kit (United States Biochemical, Cleveland, OH). Sequencing primers for both strands were designed at each 400 base-pair interval. The six newly determined IRBP intron 1 sequences (those for Cca, Cgo, Cpy, Lro, Smi, and Sbi after the code for species names in Table 1) have been deposited in GenBank under Accession Nos. CCU19748, CGU19749, CPU19750, LRU19751, SMU-19752, and SBU19753, respectively.

### Sequence alignments and phylogenetic analysis

Using the ESEE200c sequence editor (Cabot and Beckenbach, 1989), we constructed a common alignment for: (i) IRBP intron 1 sequences from 21 ceboids and human; (ii) the IRBP intron 1 and whole  $\epsilon$ -globin gene sequences from 19 ceboids and human; (iii) the IRBP and  $\epsilon$  noncoding (introns and 3' noncoding region) sequences from the same 19 ceboids and human. We inserted gaps only when they greatly increased sequence correspondence in the ungapped portions of the alignment, i.e., when they minimized the numbers of nucleotide substitutions and indels (insertions/deletions) needed to account for the descent of the aligned sequences. For phylogenetic analysis by the parsimony method, gaps were so coded that each indel regardless of its length was equivalent to a single nucleotide substitution. The alignment of the 22 intron 1 IRBP sequences (see Fig. 2) extends over 1,843 nucleotide positions, of which 350 are informative for nucleotide substitutions (i.e., at each informative position, at least two sequences share the same nucleotide and at least two other sequences share a

different nucleotide). The alignment also contains 34 informative indels (i.e., each informative indel or shared gap occurs in at least two sequences and at least two other sequences lack this indel).

Pairwise divergence values were estimated by the Jukes and Cantor (1969) method and used to compute a putative phylogenetic tree by the neighbor-joining method (Saitou and Nei, 1987) for both IRBP sequences and tandemly aligned IRBP and  $\epsilon$  noncoding (introns and 3' noncoding region) sequences.

A more thorough phylogenetic analysis, that directly utilizes the character state information in the nucleotide sequences, was preformed with the help of maximum parsimony programs (written by John Czelusniak) using the sequence data sets of IRBP alone and tandemly combined with the whole  $\epsilon$  gene. These parsimony programs were: PTRFC and SOG. PTRFC performs branch swapping with global rearrangements to search for the most parsimonious or lowest-length trees. SOG examines and determines the nucleotide substitution (NS) scores for all possible unrooted trees for a specified number of terminal branches consisting of either exterior nodes (contemporary sequences) or interior nodes (subtrees) or both sets of nodes. The option of delimiting a cutoff score, as in other branch and bound methods (Hendy and Penny, 1982), greatly speeded up the SOG runs. In the search of all possible trees it identifies and eliminates those pathways in which every tree would have a score above the cutoff score; the NS scores of only those individual trees up to an including the cutoff score were then calculated. SOG then orders the trees according to increasing NS scores and thereby obtains the strength of grouping results, i.e., the minimal number of extra mutations required to break up the clades formed by the terminal branches at the most parsimonious score. The most parsimonious trees found by our PTRFC and SOG programs not only were the same as those found by the DNAPARS program of Felsenstein (1989) but also had the same NS scores. SOG is available via anonymous ftp at the Indiana (ftp.bio.indiana.edu) software distribution site.

The strength of grouping number for an interior node in the maximum parsimony tree should not be confused with the number of putative synapomorphies (shared derived substitutions) at that interior node. If there were no homoplasy (no parallel, convergent, or back substitutions), then the strength of grouping number and number of putative synapomorphies would always be identical. However, depending on the distribution of homoplasy among the sequences, the strength of grouping number for an interior node is sometimes less than the node's number of putative synapomorphies. Consider a tree with three in group sequences (A,B, and C) and one outgroup sequence (D): if 10 putative synapomorphies support grouping A with B but 3 putative synapomorphies support grouping A with C and 2 putative synapomorphies support grouping B with C, then the node for grouping A with B in the maximum parsimony tree is supported by 10 putative synapomorphies but by a strength of grouping number of 7 [the difference between  $23 (2 \times 10 + 3)$  for the (A,C) group and  $16 (10 + 2 \times 3)$  for the (A,B) group, i.e., the minimal number of extra substitutions required for a tree that does not have A group with B]. Such strength of grouping numbers for the interior nodes of a maximum parsimony tree (or the consensus of trees with the maximum parsimony score) were first used by Goodman et al. (1982, 1985) and are equivalent to the decay index of Bremner (1988).

Bootstrap analyses were made using SEQBOOT, DNAPARS and CONSENSE programs included in the new 3.5c unpublished version of the PHYLIP package (Felsenstein, 1989), running under UNIX operational system. We carried out 2,000 bootstrap replications with 10 shufflings per each replication. That is, to find the most parsimonious tree for each sequence data set created by a bootstrap replication, the order in which the sequences were presented to the DNAPARS program were shuffled (i.e., randomly varied) 10 different times. DNAPARS constructs an initial tree and then does global branch swapping on it. The initial tree is constructed by a quite simple method that depends only on the order by which the sequences are presented to the

program. Each sequence is joined in turn to a growing tree where it adds the least length. When the last sequence is so joined, global branch swaps are then carried out in a further search for the tree with the lowest score. To compensate for the arbitrary nature of this starting tree, DNAPARS allows a number of starting trees to be used for global branch swapping. These trees are constructed by varying randomly (shuffling) the order by which the sequences are presented. The best trees (those with lowest scores) found by these swappings with different starting trees are recorded. Clearly, with different starting trees the chance of finding the true most parsimonious tree is increased. We interpret the bootstrap frequencies as alternatives to the parsimony strength of grouping numbers.

### Statistical model approach to test proposed phylogenetic groupings

Our approach, suggested by earlier work of Cavalli-Sforza and Edwards (1967) and Navidi et al (1991), is described in detail elsewhere (Czelusniak and Goodman, 1995). In brief, each stochastic evolutionary model that we test consists of an assumed stochastic process of nucleotide substitutions and a tree upon which the substitutions occur. Such a model predicts a multinomial distribution of nucleotide character patterns in the contemporary sequences. The distribution of patterns expected under each model is compared to the actual observed distribution of patterns by a goodness of fit statistic such as the loglikelihood ratio  $G^2$  after the numerical parameters for the model have been chosen to minimize the respective statistic, i.e., maximize the goodness (closeness) of fit between expected (predicted) and observed character patterns.

A nucleotide character state pattern for a set of aligned homologous sequences consists of an array of nucleotides drawn from the four nucleotides A,C,G, and T. An homologous nucleotide position where all aligned sequences are represented by a nucleotide is an instance of a nucleotide character state pattern. Binary character state patterns are obtained by converting the four nucleotides in the aligned sequences into purines

(R = A,G) and pyrimidines (Y = C,T). To obtain the observed multinomial distribution of character state patterns, we count for each different pattern the number of positions in the aligned sequences that show that pattern. For a data set of five aligned homologous sequences, there are  $4^5 = 1,024$  possible nucleotide character state patterns (there being four nucleotide possibilities in five sequences) and  $2^5 = 32$  binary character state patterns.

In the present study, as in Czelusniak and Goodman (1995), the stochastic evolutionary models used to calculate the expected multinomial distribution of nucleotide character state patterns varied in branching pattern but always employed the stochastic process of nucleotide substitutions described in Barry and Hartigan (1987). In this stochastic process, three parameters give the frequencies of the four nucleotides at the root of the tree and 12 parameters on each branch determine the 16 probabilities of one of the four nucleotides changing to another nucleotide. A simpler stochastic process is used to calculate the expected multinomial distribution of binary (R,Y) character state patterns. In this simpler process, one parameter determines the frequency of each of the two states at the root and two parameters on each branch determine the four probabilities of each of the two states remaining the same or changing to the other.

The parameters of the evolutionary models are chosen so as to minimize the loglikelihood ratio  $G^2$  goodness of fit statistic. The loglikelihood ratio  $G^2$  is defined as  $-2(\sum \eta_i \ln(P_i^{exp}(\theta)) - \sum \eta_i \ln(P_i^{obs}))$ , where  $P_i^{exp}(\theta)$ ,  $P_i^{obs}$  and  $\eta_i$  are, respectively, the expected fractions, observed fractions, and counts of nucleotide character pattern  $i$ . The summation is taken over all observed character patterns. The TOLMIN numerical minimization algorithm of Powell (1989a,b) is used to estimate the vector of parameters  $\theta$ .

For the stochastic evolutionary models that employ the same stochastic process of sequence change and vary only in the tree branching arrangement, the best branching arrangement is taken as the one that gives the lowest score for loglikelihood ratio  $G^2$  (i.e., the closest fit between the observed and expected multinomial distributions of char-

acter state patterns). However, such a "best" branching arrangement is only taken as strong evidence for that phylogenetic hypothesis if it can not be rejected on statistical grounds. Our procedure of significance testing involves numerical calculation of the distribution of the loglikelihood ratio  $G^2$  by Monte Carlo replication. Each Monte Carlo replicate draws character state patterns in proportion to their expected frequencies under the phylogenetic hypothesis. This drawing is carried out  $N$  times with  $N$  equaling the number of aligned positions having nucleotides for all species. The counts of different character state patterns obtained in this drawing from expected counts are treated now as if there were observed counts. Next, on the same tree branching arrangement (used previously for this phylogenetic hypothesis to obtain the original expected counts from the original observed counts) and with these so called "observed" counts from the Monte Carlo drawing, parameters are chosen by the TOLMIN algorithm in order to calculate a new set of expected counts with a loglikelihood value that minimizes the loglikelihood ratio goodness of fit statistic  $G^2$ . We then note for each Monte Carlo replicate if its loglikelihood ratio  $G^2$  is less than or greater than the original loglikelihood ratio  $G^2$  for that phylogenetic hypothesis. Performing a sufficient number of replicates (100–1,000, depending on the computer time needed) allows one to calculate the probability  $P$  of getting a value of the goodness of fit statistic greater than the original observed value for this statistic. In general, there is a better chance of obtaining a  $P$  value greater than 0.05 for the phylogenetic hypothesis that yielded the lowest loglikelihood ratio  $G^2$  score for the comparison of original expected counts to original observed counts of character state patterns than for the phylogenetic hypotheses that yielded higher loglikelihood ratio  $G^2$  scores. If such a  $P$  value can be obtained, it is taken as statistical support for that phylogenetic hypothesis. Conversely, a very small  $P$  value ( $<0.01$ ) for any competing phylogenetic hypothesis may be viewed as a statistical rejection of that competing hypothesis.

## RESULTS

### Phylogenetic inferences from IRBP sequences

The maximum parsimony tree for the 22 aligned IRBP intron 1 sequences (Fig. 2) is shown in Figure 3. The numbers above the lines to interior nodes represent bootstrap values (as percentages) while numbers below these lines represent the SOG parsimony program strength of grouping estimates. The NS score of the tree was 999. The cladistic results were similar to those obtained previously for the  $\epsilon$  gene (Harada et al., 1995).

Species of the same genus always strongly grouped together as  $\epsilon$  sequences (Fig. 3 in Harada et al., 1995) demonstrated for *Cebus*, *Saimiri*, *Callicebus*, *Ateles*, and *Alouatta*, and as IRBP sequences (Fig. 3, this study) demonstrate again for *Cebus*, *Saimiri*, and *Callicebus* and in addition for *Saguinus* and *Aotus*. The IRBP tree (Fig. 3) supports these intrageneric groupings with bootstrap values of 100% and strength of grouping values ranging from 25 for the *Saguinus bicolor*-*S. midas* group to 64 for the *Saimiri boliviensis*-*S. sciureus* group. These high strength of grouping values agree with the high numbers of synapomorphies yielding the intrageneric groupings. For example, as the aligned sequences (Fig. 2) reveal, the number of synapomorphies grouping *Saguinus bicolor* and *S. midas* is 25 (24 nucleotide substitutions at positions 33, 54, 89, 312, 346, 364, 380, 470, 511, 535, 647, 662, 732, 823, 908, 1212, 1229, 1306, 1307, 1394, 1413, 1510, 1525, and 1534 plus a two base deletion at 733–34).

The cladistic results with IRBP sequences (this study), like those with  $\epsilon$  sequences (Harada et al., 1995), group the 16 genera into the following monophyletic taxa: *Callithrix* and *Cebuella* into Callitrichina; Callitrichina, *Callimico*, *Leontopithecus*, and *Saguinus* into Callitrichinae; Callitrichinae, *Saimiri*, *Cebus*, and *Aotus* into Cebidae; *Cacajao* and *Chiropotes* into Chiropotina; Chiropotina and *Pithecia* into Pitheciini; Pitheciini and *Callicebus* into Pitheciinae; and *Lagothrix*, *Brachyteles*, *Ateles*, and *Alouatta* into Atelinae. The IRBP tree supports all these groupings with 99% to 100% bootstrap



```

CJA      GTGAGACCCAAGGGAGACGTGGCCAGGCCAGTCCCGGGAGTGACTTGACCCATCGTTTG 60
CPY      .....A.....
CGO      .....G.....T.....
LRO      .....G.....TG.....
SBI      .....C.....G.....T.C.....
SMI      .....G.....C.....C.....G.....T.C.....
SSC      .....C.....G.....C.....--.....G.....TG..A.....
SBO      .....C.....G.....C.....G.....TG..A.....
CKA      .....C.....GA.....G.....TG.....
CNI      .....C.....GA.....G.....TG.....
AAZ      .....C.....G.....T.....G.....TG.....
ANA      .....C.....G.....T.....G.....TG.T.....
CTO      .....C.....TG.....-----TG.....C.....
CMO      .....C.....TG.....-----TG.....C.....
CCA      .....A.....C.....G.....G.....TG..A.C.....
CSA      .....A.....C.....G.....G.....TG..A.C.....
PIR      .....C.....G.....T.....G.....G..A.C.....
LLA      .....C.....G.....T.....G.....TG.....C.....
BAR      .....C.....G.....T.....G.....TG.....C.....
ABM      .....C.....G.....T.....G.....TG.....C.....
ABE      .....A.....C.....G.....T.....A..G.....TG..A.C.....
HSA      .....A.....C.....TGAA.....G.....T..CC.....

```

```

CJA      CACATGCAGGACTCTGTGCACAGTGCATGACAATGGCTTTTAGATTGTTCTCATGTTTA 120
CPY      .....T.....
CGO      .....G.....
LRO      .....GA.....
SBI      .....G.....A.....
SMI      .....G.....A.....
SSC      .....G.....C.....
SBO      .....G.....C.....
CKA      .....G.....
CNI      .....G.....
AAZ      .....G.....A..G.....C.....
ANA      .....G.....A..G.....C.....
CTO      .....G..T.....TG.....C.....
CMO      .....G..T.....TG.....C.....
CCA      .....G.....G.....C.....G.....
CSA      .....G.....G.....G.....C.....G.....
PIR      .....G.....G.....C.....
LLA      .....G.....G.....G.....
BAR      .....GT.....G.....T.....A.....
ABM      .....G.....T.....
ABE      .....G.....T.....
HSA      .....G.....G.....C.....

```

```

CJA      AGTTTGTGGCCAGTTGAGTCCTTTCTCTTTCT-CACCCTGTTCCATCCACTCTCTGGGAC 180
CPY      .....-.....T.....
CGO      .....-.....C.....
LRO      .....-.....C.....
SBI      .....-.....
SMI      .....-.....
SSC      .....CA.....C..T.....-.....C.....
SBO      .....A.....-.....C.....
CKA      .....C.....-.....
CNI      .....C.....-.....
AAZ      .....G.....CATC.A.....
ANA      .....G.....CATC.A.....
CTO      .....-.....
CMO      .....-.....
CCA      .....G.....-.....T.....G.....
CSA      .....-.....G.....
PIR      .....T.....-.....
LLA      G.....C.....-.....AG.....
BAR      G.....C.....-.....A.....
ABM      G.....-.....T.....A.....
ABE      G.....-.....G.....A.....A.....
HSA      .....A..G..CA.....-.....A.....T.....

```

Figure 2. For legend see page 171.

CJA	CCTGGT	GCTGCTGTAGAACCTCTGTAGAACATTTCATGTTAGGTTGGTGTGAAAGTACTTT	240
CPY	.....	.....	
CGO	.....	.....	
LRO	.....	.....C.....	
SBI	.....	.....	
SMI	.....	.....	
SSC	.....	C.....G.....G.....C.....C.....	
SBO	.....	C.....G.....G.....C.....C.....	
CKA	.....	.....C.....	
CNI	.....	.....C.....	
AAZ	.....	.....C.....	
ANA	.....	.....C.....	
CTO	.....	A.....C.....C.....	
CMO	.....	A.....C.....C.....A.....G.....	
CCA	.....	.....C.....C.....A.....	
CSA	.....	.....C.....G.....CA.....	
PIR	.....	.....C.....C.....	
LLA	.....	.....C.....A.....	
BAR	.....	.....	
ABM	.....	.....T.....	
ABE	.....	A.....G.....CA.....	
HSA	.....	G.T.....-AC.....A.....CA.....	
CJA	CAATGG	CAAAACCCACAATTAATTTTGCATCAATCTCACGGGAACCCAGTTT-GAAAGCC	300
CPY	.....	.....C.....A.....G.....G.....	
CGO	.....	.....C.....A.....G.....-	
LRO	.....	.....C.....C.....------	
SBI	.....	.....C.....G.....-	
SMI	.....	.....C.....G.....-	
SSC	A.....	.....C.....G.....-	
SBO	A.....	.....C.....G.....-	
CKA	.....	.....C.....C.....G.....-	
CNI	.....	.....C.....C.....G.....-	
AAZ	T..T.....	.....C.....C.....------	
ANA	.....	T.....C.....------	
CTO	.....	A.....C.....C.....G.....-	
CMO	.....	A.....C.....C.....G.....-	
CCA	.....	.....C.C.....C.....T.....G.....-	
CSA	.....	.....C.....C.....G.....-	
PIR	.....	A.....C.....C.....G.....-	
LLA	.....	.....C.....C.....T.....G.....-	
BAR	.....	G.....C.....C.....G.....-	
ABM	.....	.....G.C.....C.....C.....T.....G.....-	A..
ABE	.....	.....C.....C.....G.....-	
HSA	.....	.....G.....C.....C.C.C.....A.....G.....-	
CJA	AAGCAATATTT	CACAGGAAGCCAGTTTGGGAAGCCCCTGGGATAGACGGAGTTTCAGCCTTG	360
CPY	.....	.....G.....T.....	
CGO	.....	.....	
LRO	-----	.....	
SBI	.....	T.....A.....	
SMI	.....	T.....A.....	
SSC	.....	G..C.....	
SBO	.....	G..C.....	
CKA	.....	C.....A.....	
CNI	.....	C.....A.....	
AAZ	-----	G.....T.....A.....	
ANA	-----	G.....-.....A.....	
CTO	.....	C.....C.....T.....C.....	
CMO	.....	C.....C.....C.....	
CCA	.....	C.....C.....	
CSA	.....	C.....C.....T.....	
PIR	.....	C.....C.....T.....	
LLA	.....	C.....C.....A.....	
BAR	.....	C.....C.....G.....A.....TG.....	
ABM	.....	C..G..C.....A.....	
ABE	.....	C.....C.....A.....A.....	
HSA	.....	C.....C.....C.....T.....A..CA.....	

Figure 2 (Continued).

CJA	GCTGGGTGGAAAGCGAGCATTGGCAGGGCGTCTCATCGTCAGTGTGGG--AGGAGGCCAA	420
CPY	.....A.....	
CGO	.....G.....G.....TG.....TT.....A.....-G.A.....	
LRO	.....G.....G.....G.....T.....A--.....-G.A.....	
SBI	..A.....G.....GC.....G.....T.....-G.A.....	
SMI	..A.....G.....GC.....G.....T.....-G.A.....	
SSC	.....G.T.....GC.....G.....T.....A.....-G.A.....G	
SBO	.....G.T.....GC.....G.....T.....A.A.....-G.A.....G	
CKA	.....G.....G.....A.T.....-T.A.....	
CNI	.....G.....G.....A.T.....-T.A.....	
AAZ	.....G.T.....G.....G.....T.....AG.A.....	
ANA	.....G.T.....G.....G.....T.....-G.A.....	
CTO	.....G.....G.....C.....A.....A.....GG.A.....TG.	
CMO	.....G.T.....G.....G.A.A.....T.....GA.A.....TG.	
CCA	...C.....G.....G.....G.....A.....A.....GG.A.....G.	
CSA	...C.....G.....G.....G.....A.....A.....A.GG.A.....G.	
PIR	.....G.....G.....G.....A.....A.....GG.A.....G.	
LLA	.....G.....G.....C.....T.....T.....A.....C.GG.A.....G	
BAR	.....G.....G.....C.....T.....T.....AGG.A.....	
ABM	.....G.....TG.....T.....T.....GT.A.....	
ABE	.....G.T.....TG.....T.....T.....GG.A.....	
HSA	.....G.T.....TG.....T.....T.....CTA.GG.A.....	
CJA	CACGTGGCAGAGGCGG-CTGTGGGCTTCACCG--CGTGCCCCACTGCAGGCCGAGAGCTC	480
CPY	.....A-.....A.....A--A.....	
CGO	..T.....T.A.....	
LRO	A.....A-.....G.....	
SBI	.....C.....A.....	
SMI	.....C.....A.....	
SSC	.....C.....	
SBO	.....T.C.....	
CKA	.....C.....C.....T.....	
CNI	.....C.....C.....T.....	
AAZ	..T.....T.-G.....C.....	
ANA	.....T.....T.-G.....C.....	
CTO	..TC.....T.T.-CA.....-A.....C.G.....	
CMO	..TC.A.....T.T.-CA.....A-A.....C.G.....T.....	
CCA	..C.....T.-A.....-A.....C.G.....	
CSA	..C.....T.-A.....-A.....C.G.....	
PIR	..C.....G.-A.....A-A.....C.G.....	
LLA	.....T.-.....CAT.....-C.G.....A.....	
BAR	.....TA.....T--T.....-C.G.....A.....	
ABM	.....A.T.-C.....TT--T.....-C.G.....A.....	
ABE	.....CC.....T.C.G.....TGT.....-C.G.....A.....	
HSA	..C.T.....T.C.AA.....A.....A-A.....-AC.T.....	
CJA	CACCCAGGCAGCACCAACTCC---ACACTGTCTCCGACCTGTCTGCTTTGCTGCATGGTC	540
CPY	.....T.....	
CGO	.....C.....T.....A.....	
LRO	.....C.A-----G.T.....T.....C.	
SBI	.....C-----A.....T.....G.....G.....	
SMI	..G.....C.....A.....T.....G.....G.....	
SSC	TG.....TC-----G.....G.T.....	
SBO	TG.....TC-----G.....C.G.T.....	
CKA	..G.....TC-----G.C.....T.....G.....T.....	
CNI	..G.....TC-----G.C.....T.....G.....T.....	
AAZ	..G.G.....TC-----G.....T.....G.....T.....	
ANA	..G.TG.....TC-----G.....T.....G.....T.....	
CTO	..G.....TC.....TCC.G.....T.....G.....T.....	
CMO	..G.....TC.....TCC.G.....C.....T.....G.....T.....	
CCA	..G.....TC.....TCC.C.....C.....T.....G.....T.....	
CSA	..G.....TC.....TCC.C.....C.....T.....G.....T.....	
PIR	..G.....TC.....TCCG.G.C.C.....T.....G.....T.....	
LLA	TG.....TC.....TCC.G.....C.....T.....G.....T.....	
BAR	TG.....TC.....TCC.G.....-T.....G.....T.....	
ABM	TG.....A.TC.....TCC.TG.....T.T.....G.....T.....	
ABE	TG.....T.....TC.....TCC.G.....C.....T.....G.....	
HSA	.....T.....TC.....A.TCC-----G.-C.....C.....	

Figure 2 (Continued).

CJA	ACAGTTGGGCAGGGCAGCATGTGTCATG-AATTCTTGCAAGGAGGGTCTGAGACCAGGG	600
CPY	.....C.....	
CGO	.....C.....A.....	
LRO	.....A.....C.....	
SBI	.....-.....	
SMI	...G.....C.....	
SSC	G.....A.....C.....	
SBO	G.....A.....C.....	
CKA	.....C.....	
CNI	.....C.....G.....	
AAZ	.....C.....TG...	
ANA	.....C.....TG...	
CTO	.....G.....-G..C.....	
CMO	.....G.....-G..C.....	
CCA	.....A.....-G..C.....A.....	
CSA	.....A.....-G..C.....	
PIR	.....A.....-G..C.....	
LLA	.....C.-G..CAT.....	
BAR	...C.....C.-G..CA.....	
ABM	.....-G..CG.....	
ABE	.....A.-G..G.....	
HSA	...A.....T..C....GG..C.....	
CJA	-TTGGGTGCAGGTAGTTTGTCTGGGAAGTGGTTACTGAAGCAGGTGTAAGGAAGGAGCA	660
CPY	-..C.....C.....	
CGO	-.....C.....G.....	
LRO	-.....C.....C.....G.....	
SBI	-.....C.....G.....G.....	
SMI	T.....C.....G.....A..G..G.....	
SSC	-.....C.....G.C..G.....G.....	
SBO	-.....C.....G.C..G.....G.....	
CKA	-.....C.....G.....G.....G..C..G	
CNI	-.....C.....G.....G.....G..C..G	
AAZ	-.....C.....G.....G.....G.....	
ANA	-.....C.....G.....G.....G.....	
CTO	-.....C.....G.....G.....G.....	
CMO	-.....C.....G.....G.....A.....G.....	
CCA	-..A.....C.A.....G.....A.....G.....	
CSA	-..A.....C.A.....G.....AG.....G...T..	
PIR	-..A.....C.A.....G.....G.....G.....	
LLA	C.....C.....G.....G.CA.....G.....	
BAR	C.....C.....G.....G.....G.....	
ABM	C.....C.....G.....G.....G.....	
ABE	C.....C.....G.....G.....G.....	
HSA	C.....CGC.....G.C.....CA.....G.....	
CJA	GGGAGAGTGAGATAGGAAGGTGACAGGCAGGTCTCTCAAAGCTGTTCTG-CTGAAGCCAG	720
CPY	.....-.....	
CGO	.....C.....C.....	
LRO	.....C.....C.....	
SBI	.A.....C.....	
SMI	.A.....C.....	
SSC	.....G.....C.....C.....	
SBO	.....C.....C.....	
CKA	.....C.....	
CNI	.....C.....	
AAZ	.....C.....	
ANA	.....C.....	
CTO	.....A.T.....C.....	
CMO	.....C.....A.T.....C.....	
CCA	.....A.....C.....	
CSA	.....A.....C.....A.....A.....	
PIR	.....A.....C.....	
LLA	.....A.....C.....	
BAR	.....A.....A.....C.....	
ABM	.....G.....	
ABE	.....A.....C.....	
HSA	.....G.T.G.....C.....GC.....	

Figure 2 (Continued).

```

CJA      GACGCTGACAAGTGTGGGGATGCTCCCAGGCACAGCCCTCTGCAGGC-AGGCCCCAGGGC 780
CPY      ...C...G.G...
CGO      ...A...G...
LRO      ...G...
SBI      ...A-...
SMI      ...A-...
SSC      ...G.A...T...T-G...
SBO      ...A...T...T-G...
CKA      ...
CNI      ...
AAZ      ...TT...G...-G...
ANA      ...TT...-G...
CTO      ...A...-GA...
CMO      ...A...-GA...
CCA      T...T...A...-GA...A...
CSA      T...T...A...-GA...
PIR      ...T...A...-GA...
LLA      .G...A...-G...
BAR      .G...-G.A...T
ABM      .GT...-G...
ABE      .G...T...CG...
HSA      ...C...G...CA...T...AG...AGT...T

```

```

CJA      TCCTGTCCACATTGGTCAAGAGTTGCCCTGAGGACATACATAACTCGGGGTTGGGCAGG 840
CPY      ...C...T...C...G...
CGO      ...G...C...TG...A...A...
LRO      ...G.T...C...G...GC...
SBI      ...G...C...G...T...G...
SMI      ...G...C...G...T...G...
SSC      ...C...TG.GC...C...TG----C...AG...
SBO      ...C...TG.C...C...TG----A...AG...
CKA      ...G...C...----G...
CNI      ...G...C...----G...
AAZ      ...G.T...C...----G...
ANA      ...G.T...C...----G...
CTO      .A...G...C...G----A...G...
CMO      .A...G...C...G----AA...G...
CCA      .A..A...G...C...C...A...G----G...
CSA      .A...G...C...C...A...G----G...
PIR      .A...T.G...C...A.TG----G...A
LLA      ...A.GG...A.C...-G----G...
BAR      ...G...A.C...-G----G...
ABM      ...G...A.C...G...A...G----G...
ABE      ...G...A.C...-G----T.TA.G...
HSA      ...G...C...-GC...A...G...

```

```

CJA      CTCCCCCTCTCTTGGAGAAGGCCTGAGCTGAGGGTGGACAGACAGGATGGTGTGTGGGAG 900
CPY      ...C...
CGO      ...A...A...T...T...
LRO      ...A...A...T...T...
SBI      ...A...
SMI      ...A...
SSC      .G.-...A...
SBO      .G.-...A...
CKA      ...A...A.G...C...
CNI      ...A...A.G...C...
AAZ      ...A...
ANA      ...A...
CTO      ...T...A...A...C...
CMO      ...T...A...A...T...C...
CCA      ...A...A...
CSA      ...A...A...A...
PIR      ...A...A...A...
LLA      ...A...T...C...G...
BAR      ...A...
ABM      ...A...G...
ABE      ...A...A...C...
HSA      ...TCT...AC...A.G...GCT...A.A...

```

Figure 2 (Continued).

CJA	AGCCTGTCAGTGGGGCCAGGTGCAGCTGAAATCAGAGGGGGCCGAGAGTGCCAACAGCAT	960
CPY	G.....	
CGO	G.....T.....-	
LRO	.....C.....T.....	
SBI	.....T.....	
SMI	.....T.....G.....	
SSC	.....C.....T.C.....	
SBO	.....A.....C.....T.C.....	
CKA	.....T.....A	
CNI	.....T.....A	
AAZ	.....T.....G.....	
ANA	.....T.....G.....	
CTO	.....C..C.....C.....T.....	
CMO	.....C..C.....CA.....T.....	
CCA	.....C.....C.....T.....	
CSA	.....C.....C.....T.....	
PIR	.....AC.....T.....A.....AAT.....	
LLA	.....A..C.....T.....A.....	
BAR	.....AA..C.....A..T.....	
ABM	.....T.....T.....A.....	
ABE	G.....C.....A..T.....T.....	
HSA	.....T.G.....G.....T.....A.....	
CJA	CTGCTACAGAATTCTCATCCCCATTTTGCATAACTGAGGCCAGAGAGGTGCAGACGGGA	1020
CPY	.....	
CGO	.....T.....G.....	
LRO	.....T.....	
SBI	.....C.....	
SMI	T.....	
SSC	.....C.....G.....A.....G.....	
SBO	.....C.....G.....A.....G.....	
CKA	.....G.....A..T..G.....G..CC.	
CNI	.....G.....A..T..G.....G..CC.	
AAZ	.....T.....G.....	
ANA	.....T.....C.....CC.....G.....	
CTO	.....T.....C.....CC.....G.....	
CMO	.....G.....T.....C.....CC.....G.....	
CCA	.....T.....T.....C.....T.....G.....	
CSA	.....T.....T.....C.....C.....G.....	
PIR	.....T.....T..G.....C.....C.....G.....	
LLA	.....T.....T.....--.....A.....C..G.....	
BAR	.....T.....T.....A.....G.....	
ABM	.....T.....A.....G.....	
ABE	.....C.....T.....A.....G.....	
HSA	.....A.....G..T.....C..GC.....G.....A.....G.....	
CJA	GTGGCCTGGACCCAGAGAGCTGTGACTGAAGGCACAGCAGGGCCTGGAGGGCAATGTCTC	1080
CPY	.....G.....G.....	
CGO	.....G.....A.....G.....	
LRO	.....G.....G.....	
SBI	.....G.....A.....G.....	
SMI	.....G.....G.....	
SSC	.....-..G.....C.....GG-.....G.....	
SBO	.....-..G.....C.....GG-.....G.....	
CKA	.....-..G..T-.....GG.....T.....G.....T	
CNI	.....-..G..T-.....GG.....T.....G.....T	
AAZ	.....G.....GG.....G.....	
ANA	.....G.....GG.....G.....	
CTO	.....GT.....A..G.....GG.....A.....G.....	
CMO	.....C.....GT.....ACG.....GG.....A.....A.....G.....	
CCA	.....G.....A.....G.....GG.....T.....G.....	
CSA	.....G.....A.....G.....GG.....G.....	
PIR	.....A.....G.....A.....G.....GG.....G.....	
LLA	.....GT.....G.....GG.....G.....	
BAR	.....GT.....G.....GG.....G.....	
ABM	.....GT.....G.....G.....G.....	
ABE	.....GT.....C.....G.....GG.....GC.....	
HSA	.....AGG.....G.....GG.....G.....	

Figure 2 (Continued).

CJA	T-GTCAGCACAGGCTCCTTGCCCCAGTCCAGCTCACCAAGTCCTGCTGCCCTCCTGCAGC	1140
CPY	..-A.....	
CGO	..-.....G.....CA.....C.T..	
LRO	..-.....A.....C.....	
SBI	..-.....C.....C.....	
SMI	..-.....C.....C.....	
SSC	..-C.....T.....C.C.....C.....	
SBO	..-C.....T.....C.C.....C.....	
CKA	..-.....C.....C.....C.....	
CNI	..-.....C.....C.....C.....	
AAZ	..-.....C.....C.....	
ANA	..-.....C.....C.....	
CTO	..-.....T.....C.....C.....	
CMO	..-.....T.....G.....	
CCA	..-.....T.....	
CSA	..-.....C.....C.....	
PIR	..-.....T.....C.....C.....	
LLA	..T.....	
BAR	..T.....G.....	
ABM	..T.....G.....	
ABE	..---T.....C.....	
HSA	..-...ATGA..T.....--..T...A.T.....A.....CT.....A.....	
CJA	CTTAGAGAGGGAGGAGGAGGTGCACCCACGTGAAAGTAGCCTGTGCTAGGCTTTCAGAAT	1200
CPY	..C.....T..T..G.....	
CGO	..C.....T.....G.....	
LRO	..C.....T.....G.....	
SBI	..C.....T.....G.....	
SMI	..C.....T.....G.....	
SSC	..C.....C..G..C..G.....	
SBO	..C.....C..G..C..G.....	
CKA	..C.....TG..C..G.....	
CNI	..C.....T..C..G.....	
AAZ	..A.....T...A..G.....	
ANA	..A.....T...A..G.....	
CTO	..C.....CT...G.....	
CMO	..C.....CT...G.....	
CCA	..AT.....CTT...A..G.....C.....	
CSA	..AT.....CTT...A..G.....C.....	
PIR	..C.....CT...A..G.....	
LLA	..C.....T...G.....	
BAR	..T.....AG.....	
ABM	..TT...A..G.....	
ABE	..T.....	
HSA	..T.....CG.....G.....	
CJA	CCCCAGTTTGCAAATTAAATTGCT----CCTTCCTTTCTGGTATAGCCAAGGTTTACAATT	1260
CPY	..A.....C.....	
CGO	..A.....C.....	
LRO	..C.....C.....	
SBI	..C.T.....G.....	
SMI	..C.T..C.....G.....	
SSC	..C.....G.....T.....	
SBO	..C.....G.....T.....	
CKA	..C.....G.....	
CNI	..C.....G.....	
AAZ	..A.....C.....T.....	
ANA	..A.....C.....T.....	
CTO	..C.....G..C.....	
CMO	..C.....G..C..C.....	
CCA	..C.....C.....C.....	
CSA	..C.....C.....C.....	
PIR	..C.....C..G..ATAG.....C.....	
LLA	..C..G.....A.....	
BAR	..C..G.....A.....	
ABM	..C..T.....	
ABE	..T.....C.....CC.....T.....A.....	
HSA	..C.....C.....CAT.....CA.....G.....	

Figure 2 (Continued).

CJA TAAAGTCAGATGTGGATTTCAGATTCTAGCTCCACCACTTATTGACTGT--AACCTGGGAC 1320  
CPY .....A.....G.....T.....--  
CGO .....G.....G.....--  
LRO .....G.....CA--  
SBI .....G.....CA--  
SMI .....G.....CA--  
SSC .GG.....G.....G.....C.....GT..  
SBO .GG.....G.....G.....C.....GT..  
CKA .GG.....T.....G.....GT..  
CNI .GG.....T.....G.....GT..  
AAZ .GG.....G.....G.....GT..  
ANA .GG.....G.....G.....GT..  
CTO .GG.A.....T.....G.....C.....GT..  
CMO .GG.....G.....C.....GT...C..  
CCA .GG.....C.....G.....GT..  
CSA .GG.....C.G.G.....GT..  
PIR .GG.....C.....G.....GT..T..  
LLA .GG.....G.....AT..  
BAR .GG.....C.....G.....GT..  
ABM .GG.....G.....GT..  
ABE GGG.....G.....GT..  
HSA .GG.....G.....C.....GTG.....A..

CJA TAGTTACTTAATCTCACTGTGCTTCAGTTTTTCCATGGAAAAGATGGGGATCATGTTATC 1380  
CPY .....  
CGO .....  
LRO .....G.....  
SBI .....  
SMI .G.....  
SSC .....T.....T.....  
SBO .....T.....T.....  
CKA .....-.....  
CNI .....-.....  
AAZ .....G.....C.....  
ANA .....G.....  
CTO .....  
CMO .....  
CCA .....T.....C..  
CSA .....C.....C..  
PIR .....C.....C..  
LLA .....  
BAR .....G.....  
ABM .....  
ABE .....  
HSA .....T.....G.....

CJA TCCT-TACAGGTGGCTGTGAGGATGATGATAAGCTCTACAAAGTGCTTAGTACAGAGCCA 1440  
CPY .....-.....G.....G..  
CGO .....-.....G.....G..  
LRO .....-.....C.....G..  
SBI .....-.....A.....A.....G..  
SMI .....-.....A.....A.....G..  
SSC .....G.....G.....G..  
SBO .....G.....A.....G.....G..  
CKA .T..G.....G.....ATG.....G..  
CNI .T..G.....G.....A.G.....G..  
AAZ .....G.....C.....G.....G..  
ANA .....G.....G.....G..  
CTO .....G.G.....C.....G.....G..  
CMO .....G.G.....C.....G.....G..  
CCA .....G.G.....A.....G..  
CSA .....GG.G.....A.....G..  
PIR .....G.G.....A.....G..  
LLA .....G.G.....G..  
BAR .....G.G.....G..  
ABM .....G.G.....T.....G..  
ABE .....G.G.....G..  
HSA .....G.....G..

Figure 2 (Continued).



CJA	GGCACCTGTTAAAGGTAGCTAACA-TCTTCCAATCCTGGCCCAGTGGAGGGGAAGATGAG	1500
CPY	...T.....-.....C.....	
CGO	.....-.....C.....A.....	
LRO	.....-.....C.....A.....	
SBI	.....-.....C.....	
SMI	.....-.....C.....	
SSC	.....C...A...GCT...T...C...C.....	
SBO	.....C...A...GCT...T...C...C.....	
CKA	...T.....A...G-.....C...G.....T.....	
CNI	...T.....A...G-.....C...G.....T.....	
AAZ	.....A...-.....C.....A.....	
ANA	.....A...-.....C.....A.....	
CTO	.....A...-T.....C.....	
CMO	.....A...-T.....C...G.....	
CCA	...G.....A...-C...T...C.....	
CSA	...G.....A...-C.....C.....	
PIR	...A.....A...-C.....CG.....A.....	
LLA	.....A...G-.....C.....	
BAR	.....A...-.....C.....	
ABM	.....GA.....-.....C.....C.....	
ABE	.....A...-.....C.....	
HSA	..TG...G.....A...-.....C.....C-----	
CJA	CTTAGAGACGTTGGGAAGCATCTGGCGAGGCAGGAGGAATCAGAGAGGAAACCATTTCCTG	1560
CPY	.....C.....	
CGO	.....A.....	
LRO	...T.....	
SBI	...A...T...T...A...G.....	
SMI	...A...T...T...A...G.....	
SSC	...T...T...TT...A...G...A.....	
SBO	...T...T...TT...A...G...A.....	
CKA	...T...T...ATT...G.....	
CNI	...T...T...ATT...G.....	
AAZ	...T...T...TT...C...G.....	
ANA	...T...T...TT...C...G.....	
CTO	...T...T...TT...G.....	
CMO	...T...T...TT...G.....	
CCA	...A...T...T...ATT...G.....	
CSA	...A...T...G.ATT...G...G.....	
PIR	...G.A...T...G...ATT...G.....	
LLA	..C.A...A...T...T...G.....	
BAR	..C...A...T...T...G.....	
ABM	..C...C.A...T...A.TT...G.....	
ABE	..C...A...T.C...TT...A...G.....	
HSA	...A.-A...T...TA...TT...AGG...G...G.....	
CJA	GGCCTTCCAGCTCTGAACACCAGAG-----CAGACAAGAACATCCTCTGCAAGG	1620
CPY	-----G...G.....	
CGO	-----G...G...G.T.....	
LRO	...CA.....G...G.....	
SBI	-----G...G.....	
SMI	-----G...G.....	
SSC	-----G...G.....	
SBO	-----G...G.....	
CKA	...C.....G...G...C.....	
CNI	...C.....G...G...C.....	
AAZ	-----G...G.....	
ANA	-----G...G.....	
CTO	...G...G.CATGTGGGTGG...AGG...G.....	
CMO	...G...A.G.CATATGGGTGG...AGG...G.....	
CCA	...G...G.CATGTGGGTGG...AGG...G.....	
CSA	...G...G.CATGTGGGTGG...AGG...G.....	
PIR	...G...G.CATGTGGGTGG...AGG...G.....	
LLA	...G...G.CATGTGGGTGG...GGA.G...T.....	
BAR	...G...G.CATGTGGGTGG...TGGA.G.....	
ABM	...G...G.CATGTGGGTGG...TGG...G...C.....	
ABE	...G...G.CATGTGGGTGG...G...G.....	
HSA	...T...A.G...G.CATGTGGGTGG...G...G.....	

Figure 2 (Continued).

CJA	AGGCTTCCCATGGATCACACATGTCCCAGTGGCATGTCAC-ATCCCAGACATGCCACTGG	1680
CPY	.....	
CGO	.....	
LRO	.....A.....	
SBI	.....	
SMI	.....C.....	
SSC	.....G.....	
SBO	.....G.....	
CKA	.....G.....A.....	
CNI	.....G.....A.....	
AAZ	.....	
ANA	.....	
CTO	.....G.....A.....	
CMO	.....G.....A.....	
CCA	...C.....A.....	
CSA	...C.....A.....G.....	
PIR	...C.....	
LLA	.....A.....T.....A.....	
BAR	.....C.....A.....	
ABM	.....C.....A.....T.....	
ABE	.....G.....A.....	
HSA	.....G.....A.....	
CJA	GAAAGTCCTGGGTGCCTACCGACTCCTTCAGAAATGTCAGTTCTCTGTCCCATGCCCTTA-	1740
CPY	.....	
CGO	.....T.....	
LRO	.....TT.....A.....	
SBI	.....T.....C.....C.....	
SMI	.....T.....A.....	
SSC	T.....T.....AC.....	
SBO	.....T.....AC.....	
CKA	.....T.....	
CNI	.....T.....	
AAZ	.....CA.....T.....	
ANA	.....C.....T.....	
CTO	.....C.....T.....	
CMO	.....C.....TT.....T.....	
CCA	A.....T.....A.G.....	
CSA	.....C.....T.....	
PIR	.....C.....T.....	
LLA	.G.....C.....T.....	
BAR	.....C.....T.....	
ABM	.....C.....T.....	
ABE	.G.....C.....G.....C.....	
HSA	.....CT.....T.....TA.T.....T.....	
CJA	---ATATTTCCACGACATAAAGGCGATCCGTGGCACCTGCTTTCCTGGGCTCCAAAACC	1800
CPY	---.....T.....T.....T.....	
CGO	---.....T.....	
LRO	---.....T.....G.....T.....	
SBI	---.....T.T.....A.....GT.....	
SMI	---.....G.....T.....	
SSC	---.....GT.....G.....A.....G.....T.....	
SBO	---.....GT.....G.....A.....G.....C.....T.....	
CKA	GTC.....T.....C.....A.....T.....	
CNI	GTC.....T.....C.....A.....T.....	
AAZ	---.....T.....A.....G.....	
ANA	---.....T.....A.....TG.....G.....	
CTO	---.C.....T.....C.G.....G.....T.....	
CMO	---.....T.....C.A.....T.....	
CCA	---.....T.....C.....T.....G.....T.....	
CSA	---.....T.....C.....T.....G.....T.....	
PIR	---.....T.....T.....T.....T.....	
LLA	---.....T.TG.....A.G.....T.....G.....	
BAR	---.....T.....G.....A.....TG.....T.....	
ABM	---.....T.....G.....ATG.....T.....	
ABE	---.....GT.....TG.....A.....A.G.....T.....	
HSA	---.....G.....T.....T.....	

Figure 2 (Continued).

CJA	GGCTG-CCCTCCTGACACTGAGCAGGATCTCCAACCTTTACAG	1843
CPY	.....	
CGO	.....C.....	
LRO	A.....C.....	
SBI	.....C.....	
SMI	.....C.....	
SSC	.....T.....T.....G...	
SBO	.....T.....C.....G.G.	
CKA	.....G.....C.....	
CNI	.....G.....C.....	
AAZ	.....C.....G...	
ANA	.....C.....G...	
CTO	.....C.....C.....	
CMO	.....C.....C.....	
CCA	A.....C.....G...	
CSA	A.....C.....G...	
PIR	.....C.....	
LLA	.....C.....	
BAR	.....C..T.....	
ABM	.....C.....	
ABE	.....C.....	
HSA	.....CT.....T....C.....	

Fig. 2. Aligned IRBP intron 1 sequences representing the 21 ceboid species listed in Table 1 and *Homo sapiens* (Hsa). Dot means same as the nucleotide in the top row. Dash means gap or part of a gap (if contiguous with another dash in the row). Gaps designate indels (see text).

values and relatively high strength of grouping values ranging from 11 for Cebidae to 27 for Callitrichina. Again, there is close agreement between the strength of grouping values in the IRBP tree (Fig. 3) and the numbers of synapomorphies in the aligned sequences (Fig. 2) yielding these intergeneric groupings. For example, the number of synapomorphies grouping *Callithrix jacchus* and *Cebuella pygmaea* is 28 (27 nucleotide substitution at positions 45, 52, 71, 262, 285, 372, 379, 385, 390, 412, 496, 634, 694, 790, 797, 817, 878, 1135, 1165, 1173, 1210, 1287, 1479, 1603, 1606, 1700, and 1824 plus a two base deletion at 409–10). Due to G to A substitutions, *Cebuella*, *Callimico*, and *Leontopithecus* share A at position 436 which accounts for the strength of grouping number for the *Callithrix-Cebuella* node being 27 rather than 28, i.e., the tree that breaks up this node with the minimum number of extra NS over the maximum parsimony number groups *Cebuella* with a *Callimico-Leontopithecus* group rather than with *Callithrix*.

The  $\epsilon$  (Harada et al., 1995) and present IRBP results further agree, but with lesser bootstrap and strength of grouping values, in sister-grouping *Saimiri* and *Cebus* into Cebinae, *Lagothrix* and *Brachyteles* into Brachytelina, and Brachytelina and *Ateles*

into Atelini. There are only a few cladistic disagreements between the  $\epsilon$  and IRBP results. The  $\epsilon$  tree joined a *Callimico-Callitrichina* clade to a *Leontopithecus-Saguinus* clade, whereas the IRBP tree first joins Callitrichina to a *Callimico-Leontopithecus* clade and then adds *Saguinus* to this group. The  $\epsilon$  tree joined *Aotus* to Callitrichinae before adding Cebinae (*Saimiri*, *Cebus*) whereas the IRBP tree joined Cebinae to Callitrichinae before adding *Aotus*. However, for only two extra NS the IRBP sequences can switch the positions of *Aotus* and Cebinae, i.e., join *Aotus* first to the callitrichine clade. Nevertheless with IRBP sequences it takes 11 extra NS to break up the cebid clade by removing *Aotus* from it. Finally the  $\epsilon$  tree joined the pitheciine clade to the ateline clade, whereas the IRBP tree joined the pitheciine clade to the cebid clade.

#### Phylogenetic inferences from tandemly combined IRBP and $\epsilon$ sequences

In Figure 4 we present the maximum parsimony tree obtained from the alignment of  $\epsilon$ -globin sequences combined in tandem with IRBP intron 1 sequences for 20 anthropoid species (19 ceboids and human). Two most parsimonious trees were obtained each with a score of 1,609. In one tree, *Aotus* groups

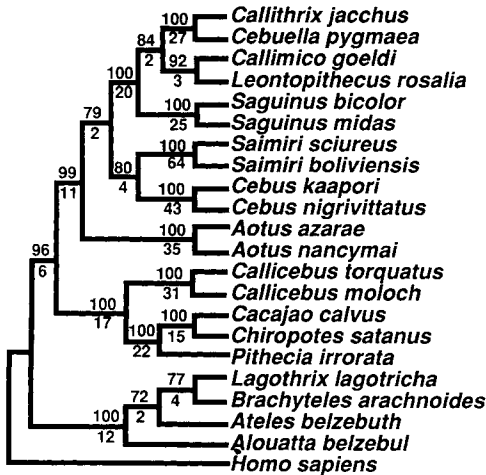


Fig. 3. Maximum parsimony tree found for the 22 aligned IRBP intron 1 nucleotide sequences. Nucleotide substitution (NS) score = 999. Numbers above lines to interior nodes represent bootstrap values (as percentages) obtained in 2,000 replications with 10 shuffles per replication. Numbers below lines to interior nodes represent parsimony strength of grouping values (i.e. at each interior node, minimum number of additional substitutions required to break up the group designated by that node).

first with the callitrichine clade and then the *Cebus-Saimiri* clade joins this *Aotus*-callitrichine clade. In the other tree, positions of *Aotus* and the *Cebus-Saimiri* clade are reversed. Thus, the consensus of the two trees depicts a trichotomous separation of *Aotus*, *Cebus-Saimiri* clade, and callitrichine clade. The arrangement produced by the tandemly combined sequences yields stronger support than either  $\epsilon$  or IRBP sequences alone for the sister grouping of *Brachyteles* with *Lagothrix*, also stronger support than IRBP sequences alone for the sister grouping of *Ateles* with the *Brachyteles-Lagothrix* clade.

Table 2 shows percent sequence divergence values derived from the tandemly combined noncoding  $\epsilon$  (introns 1 and 2 and 3' noncoding region) and IRBP intron 1 sequences. The divergence between *Callithrix* and *Cebuella* (1.6%) as well as the divergence between *Cacajao* and *Chiropotes* (2.0%) are remarkably low for intergeneric comparisons.

The Neighbor-joining tree (not shown)

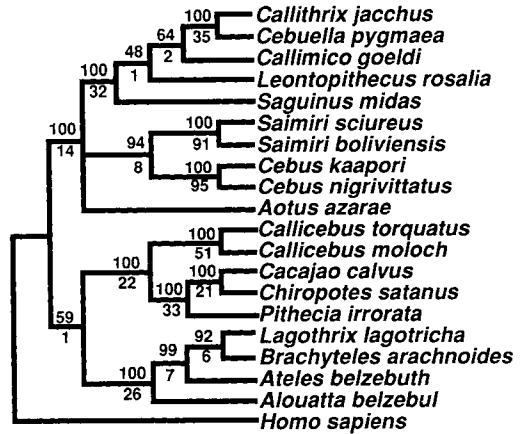


Fig. 4. Consensus maximum parsimony tree found for 20 aligned tandemly combined  $\epsilon$ -globin gene and IRBP intron 1 sequences. NS = 1,609. Numbers above lines to interior nodes represent bootstrap values (as percentages) obtained in 2,000 replications with 10 shuffles per replication. Numbers below lines to interior nodes represent parsimony strength of grouping values.

constructed from the percent sequence divergence values in Table 2 has almost the same branching arrangement as the maximum parsimony tree (Fig. 4) for combined  $\epsilon$  and intron 1 IRBP sequences. The only difference is that *Aotus* [as with  $\epsilon$  sequences along (Harada et al., 1995)] is sister of the callitrichines.

#### Testing phylogenetic inferences by stochastic evolutionary models

Table 3 presents the loglikelihood ratio  $G^2$  scores for representative series of either four taxa or five taxa trees, using the Barry and Hartigan (1987) stochastic process for nucleotides. Table 4 presents loglikelihood ratio  $G^2$  scores for either four taxa or five taxa trees, using the simpler binary (Y and R) code. It can be seen from these scores that the best fits of expected to observed character state patterns were obtained by those trees that paralleled, in their grouping of representatives of different ceboid clades, the groupings of these clades in the maximum parsimony trees. Typically, if a series involved sequences showing very little homoplasy (as was more often so for  $\epsilon$  sequences than for IRBP sequences), the tree with the lowest loglikelihood  $G^2$  score passed

TABLE 2. Sequence divergence values ( $\times 100$ ) from pairwise comparisons of 20 19 exons and 5' noncoding and IRBP-intron 1 sequences

	Hsa	Cja	Cpy	Cgo	Lro	Smr	Aaz	Cka	Cni	Ssc	Sbo	Cca	Pir	Cto	Cmo	Bar	Lla	Abm	Abe
Cja	11.9																		
Cpy	12.3	1.6																	
Cgo	11.3	3.5	3.9																
Lro	11.7	4.2	4.5	3.6															
Smr	12.0	4.1	4.5	3.7	3.9														
Aaz	11.3	5.7	6.2	5.2	5.6	5.6													
Cka	11.9	6.3	6.7	6.1	6.7	6.3	5.7												
Cni	11.9	6.3	6.7	6.1	6.7	6.3	5.7	0.2											
Ssc	12.4	7.0	7.5	6.5	7.0	6.6	6.2	6.6	6.5										
Sbo	12.5	7.1	7.6	6.4	7.0	6.6	6.3	6.6	6.5	0.8									
Cca	12.0	8.0	8.4	7.7	8.0	7.2	7.9	7.9	7.9	8.4	8.5								
Csa	11.8	8.1	8.5	7.4	7.8	7.1	7.5	7.5	7.5	8.4	8.3	2.0							
Pir	11.3	7.6	8.0	7.2	7.5	7.4	6.6	7.3	7.2	7.9	7.9	3.2	2.9						
Cto	11.6	7.7	8.0	7.2	7.7	7.5	6.7	7.2	7.1	7.9	7.8	5.6	5.1	4.9					
Cmo	12.0	8.0	8.4	7.7	8.2	8.0	7.2	7.5	7.4	8.4	8.4	6.0	5.8	5.5	2.0				
Bar	11.1	7.0	7.4	6.8	7.2	7.0	6.1	7.0	7.0	7.7	7.7	6.4	6.3	6.0	6.5				
Lla	11.5	7.1	7.7	7.0	7.2	7.2	6.4	7.2	7.1	7.8	7.8	6.6	6.7	6.4	5.9	3.2			
Abm	11.2	6.9	7.2	6.7	7.0	6.9	6.1	6.7	6.7	7.7	7.7	6.1	6.2	5.8	6.4	3.1	3.7		
Abe	11.6	7.6	8.0	7.3	8.0	7.6	6.9	7.7	7.6	8.0	8.0	7.1	7.0	6.6	7.1	4.3	4.8	4.4	

the Monte Carlo tests ( $P > 0.05$ ), whereas trees with higher scores in most instances failed ( $P < 0.01$ ). In particular, these Monte Carlo results for both the nucleotide and binary codes with  $\epsilon$  sequences, but for only the binary code with IRBP sequences, provided strong statistical evidence for *Callicebus* being the sister of pitheciines. These Monte Carlo results also provide statistical evidence for the sister grouping of *Cebus* with *Saimiri* and for the cebid clade of *Cebus*, *Saimiri*, *Aotus*, and callitrichines. The grouping of pitheciines with atelines by  $\epsilon$  sequences was supported statistically, but neither this grouping nor the two alternatives to it (pitheciines either with cebids or with a cebid-atelid ancestor) were supported statistically by IRBP intron 1 sequences. Thus in balance a sister group relationship of pitheciines with atelines still appears to be a viable hypothesis from these molecular trees.

## DISCUSSION

### Ateline clade

The cladistic arrangement obtained by both maximum parsimony and neighbor-joining methods identifies *Alouatta* as the sister group of atelins (*Ateles*, *Brachyteles*, and *Lagothrix*). *Alouatta* groups with the atelins in 100% of bootstrap replicates and 26 extra mutations are necessary to break up this ateline clade (Fig. 4). The great similarity of *Alouatta* with the other atelines is evident in the majority of classical arrangements (Cabrera, 1958; Hill, 1962; Napier and Napier, 1967; Hershkovitz, 1977). Its placement as the earliest diverging member of the ateline clade has also been corroborated by cladistic analyses based on morphological data (Rosenberger, 1981; Ford 1986). Nevertheless, some authors disagree with this view. Dunlap et al. (1985) grouped *Alouatta* with *Lagothrix* on the basis of forelimb morphology, while Kay (1990) grouped *Alouatta* and *Brachyteles* on the basis of dental characteristics. However, this latter grouping can be attributed to convergent features reflecting the folivorous diet of *Alouatta* and *Brachyteles* (Anthony and Kay, 1993).

With both IRBP and  $\epsilon$  sequences, *Brachyteles* and *Lagothrix* share the most

TABLE 3. Loglikelihood ratios  $G^2$  and estimates of  $P$  values for models using the same process of nucleotide substitution

Tree	IRBP observed loglikelihood $G^2$	$\epsilon$ observed loglikelihood $G^2$
((Ssc,Aaz),(Pir,Bar),Hsa)	185.7 (0%) <sup>1</sup>	75.2 (13%)
((Ssc,Pir),(Aaz,Bar),Hsa)	236.3 (0%)	174.3 (0%)
((Ssc,Bar),(Aaz,Pir),Hsa)	232.6 (0%)	174.6 (0%)
((Ssc,Hsa),(Pir,Bar),Aaz)	211.5 (0%)	104.7 (0%)
((Ssc,Pir),(Hsa,Bar),Aaz)	228.3 (0%)	173.8 (0%)
((Ssc,Bar),(Hsa,Pir),Aaz)	224.5 (0%)	159.8 (0%)
((Ssc,Aaz),(Hsa,Bar),Pir)	184.5 (0%)	112.5 (0%)
((Ssc,Bar),(Aaz,Hsa),Pir)	235.2 (0%)	170.3 (0%)
((Ssc,Hsa),(Aaz,Bar),Pir)	230.5 (0%)	169.5 (0%)
((Ssc,Aaz),(Pir,Hsa),Bar)	178.7 (0%)	105.8 (0%)
((Ssc,Pir),(Aaz,Hsa),Bar)	237.3 (0%)	171.5 (0%)
((Ssc,Hsa),(Pir,Aaz),Bar)	228.8 (0%)	172.0 (0%)
((Hsa,Aaz),(Pir,Bar),Ssc)	218.5 (0%)	105.2 (0%)
((Hsa,Pir),(Aaz,Bar),Ssc)	225.6 (0%)	159.2 (0%)
((Hsa,Bar),(Aaz,Pir),Ssc)	227.5 (0%)	174.9 (0%)
((Ssc,Aaz),Abm,Hsa)	79.4 (0%)	34.8 (44%)
((Ssc,Abm),Aaz,Hsa)	103.2 (0%)	56.0 (0%)
((Ssc,Hsa),Aaz,Abm)	98.8 (0%)	55.1 (0%)
((Ssc,Cka),Aaz,Cja)	80.5 (0%)	27.8 (30%)
((Ssc,Aaz),Cka,Cja)	88.6 (0%)	49.0 (0%)
((Ssc,Cja),Cka,Aaz)	77.6 (0%)	57.5 (0%)
((Csa,Cca),(Cmo,Cto),Pir)	48.0 (0%)	28.1 (14%)
((Csa,Cmo),(Cca,Cto),Pir)	583.3 (0%)	347.4 (0%)
((Csa,Cto),(Cca,Cmo),Pir)	585.2 (0%)	347.4 (0%)
((Csa,Cto),(Pir,Cmo),Cca)	563.5 (0%)	343.2 (0%)
((Csa,Pir),(Cto,Cmo),Cca)	127.7 (0%)	87.8 (0%)
((Csa,Cmo),(Cto,Pir),Cca)	567.6 (0%)	340.9 (0%)
((Csa,Cca),(Cto,Pir),Cmo)	424.1 (0%)	243.3 (0%)
((Csa,Cto),(Cca,Pir),Cmo)	578.8 (0%)	318.1 (0%)
((Csa,Pir),(Cca,Cto),Cmo)	565.5 (0%)	321.2 (0%)
((Csa,Pir),(Cca,Cmo),Cto)	565.4 (0%)	321.2 (0%)
((Csa,Cca),(Pir,Cmo),Cto)	419.6 (0%)	242.9 (0%)
((Csa,Cmo),(Pir,Cca),Cto)	577.6 (0%)	318.1 (0%)
((Cto,Cmo),(Pir,Cca),Csa)	127.7 (0%)	87.6 (0%)
((Cto,Pir),(Cmo,Cca),Csa)	562.7 (0%)	340.4 (0%)
((Cto,Cca),(Cmo,Pir),Csa)	565.3 (0%)	342.7 (0%)
((Bar,Pir),Aaz,Hsa)	90.6 (0%)	35.0 (32%)
((Bar,Aaz),Pir,Hsa)	77.4 (0%)	70.3 (0%)
((Bar,Hsa),Pir,Aaz)	87.4 (0%)	62.6 (0%)

<sup>1</sup>The percent values in parentheses are estimated  $P$  values (number of Monte Carlo replicates out of 100 having  $G^2$  greater than that of the observed).

recent ancestor and have *Ateles* as their sister group. Moreover, this arrangement is supported by cytogenetic data: *Brachyteles* and *Lagothrix* share a similar karyotype with  $2n = 62$  chromosomes while the diploid number in *Ateles* varies from  $2n = 32$  to  $2n = 34$  (Dutrillaux, 1988; Pieczarka et al., 1989). Thus, we divide the tribe Atelini into subtribes Brachytelina (*Brachyteles*, *Lagothrix*) and Atelina (*Ateles*).

These molecular and chromosomal data disagree with all previous arrangements based on morphological characters. Zingesser (1973) grouped *Ateles* first with *Lagothrix*, while Rosenberger and Strier (1989) grouped *Ateles* first with *Brachyteles*. Ford

(1986) shows an unresolved *Ateles*/*Brachyteles*/*Lagothrix* trichotomy. This controversy is probably a consequence of the apparently very short span of time separating the emergence of these three lineages. In that regard our molecular data show very small differences in nucleotide diversity among the three genera (Table 2).

#### Pitheciine clade

The cladistic arrangement obtained with molecular data has *Chiropotes* as the sister of *Cacajao*, and then *Pithecia* as the sister of the *Chiropotes*-*Cacajao* clade. *Callicebus* appears as an independent lineage closely

TABLE 4. Loglikelihood ratios  $G^2$  and estimates of  $P$  values for models using the purine, pyrimidine binary code

Tree	IRBP observed loglikelihood $G^2$	$\epsilon$ observed loglikelihood $G^2$
((Ssc,Aaz),Abm,Hsa)	1.62 (32%) <sup>1</sup>	0.52 (81%)
((Ssc,Abm),Aaz,Hsa)	5.37 (1%)	2.68 (8%)
((Ssc,Hsa),Aaz,Abm)	5.76 (1%)	2.68 (6%)
((Cto,Pir),(Aaz,Cka),Hsa)	6.68 (34%)	7.39 (19%)
((Cto,Aaz),(Pir,Cka),Hsa)	81.28 (0%)	51.61 (0%)
((Cto,Cka),(Pir,Aaz),Hsa)	78.94 (0%)	51.61 (0%)
((Hsa,Pir),(Aaz,Cka),Cto)	49.50 (0%)	32.65 (0%)
((Hsa,Aaz),(Pir,Cka),Cto)	81.28 (0%)	51.61 (0%)
((Hsa,Cka),(Aaz,Pir),Cto)	79.89 (0%)	48.90 (0%)
((Cto,Hsa),(Aaz,Cka),Pir)	49.50 (0%)	35.57 (0%)
((Cto,Aaz),(Hsa,Cka),Pir)	79.89 (0%)	48.90 (0%)
((Cto,Cka),(Aaz,Hsa),Pir)	78.94 (0%)	51.61 (0%)
((Cto,Pir),(Hsa,Cka),Aaz)	14.60 (8%)	11.43 (9%)
((Cto,Hsa),(Pir,Cka),Aaz)	81.28 (0%)	51.61 (0%)
((Cto,Cka),(Pir,Hsa),Aaz)	78.94 (0%)	48.54 (0%)
((Cto,Pir),(Aaz,Hsa),Cka)	14.84 (7%)	13.32 (5%)
((Cto,Aaz),(Pir,Hsa),Cka)	81.28 (0%)	48.55 (0%)
((Cto,Hsa),(Pir,Aaz),Cka)	81.28 (0%)	51.61 (0%)

<sup>1</sup>The percent values in parentheses are estimated  $P$  values (number of Monte Carlo replicates out of 100 having  $G^2$  greater than that of the observed).

linked to the three pitheciins. The monophyly of the pitheciins, represented by *Pithecia*, *Chiropotes*, and *Cacajao*, has been demonstrated in the majority of previous studies using morphological (Rosenberger, 1981; Ford, 1986; Kay, 1990) and biochemical markers (Schneider et al., 1993). Studies based on immunological distances and karyotypes have reached similar conclusions using only two of the three genera: *Pithecia* and *Cacajao* (Cronin and Sarich, 1975; Dutrillaux, 1988) or *Cacajao* and *Chiropotes* (Baba et al., 1980).

Most previous arrangements have identified *Chiropotes* and *Cacajao* as the most closely related lineages, but De Boer (1974) proposed an alternative arrangement based on chromosomal data, grouping *Pithecia* ( $2n = 48$ ) with *Cacajao* ( $2n = 44$  to  $46$ ), and identifying *Chiropotes* ( $2n = 54$ ) as the oldest member of the clade.

The close similarity of both *Aotus* and *Callicebus* with the pitheciins suggested by Rosenberger (1981) and Rosenberger et al. (1990) is not corroborated by our molecular data. While *Callicebus* is strongly connected to the pitheciins, *Aotus* is clearly connected to cebines (*Saimiri*, *Cebus*) and callitrichines (Figs. 3 and 4).

De Boer (1974), in a phylogenetic reconstruction based on chromosome diploid num-

ber, depicts an ancestral karyotype close to  $2n = 54$ . He places *Callicebus* ( $2n = 20$  to  $50$ ) as the basal lineage of the pitheciines, followed by *Aotus* ( $2n = 48$  to  $54$ ), *Chiropotes* ( $2n = 54$ ), and *Pithecia-Cacajao* ( $2n = 44$  to  $46$ ) as the most derived clade.

Based on our molecular data, we propose that the ancestral karyotype of pitheciines was close to  $2n = 46$  instead of  $2n = 54$ . The first lineage to emerge, *Callicebus*, preserved the primitive karyotype in *Callicebus moloch* and *C. donacophilus*, and experienced a drastic chromosome reduction in *C. torquatus*. According to this view, *C. torquatus* is the most recent species in the genus and not the oldest one as proposed by Hershkovitz (1990). According to our interpretation, *Pithecia* with the conserved chromosome number  $2n = 46$  originated from the branching event that separated it from the ancestor of *Cacajao* and *Chiropotes*. The last radiation in the pitheciine clade, involving probably drastic chromosome rearrangements, led to *Cacajao* with ancestral chromosome number  $2n = 48$  and to *Chiropotes* with  $2n = 54$ .

### Cebid clade

Cebidae, in the cladistic classification that Rosenberger et al. (1990) proposed based on morphological characters, is closely similar

to that based on DNA sequences presented here, the only difference being the allocation of *Aotus*. We place *Aotus* along with *Cebus*, *Saimiri*, and the callitrichines in the family Cebidae, whereas Rosenberger et al. (1990) has *Aotus* as the sister of *Callicebus* and groups this *Aotus-Callicebus* clade with the pitheciins. Defining *Aotus* as the sister of *Callicebus* using either  $\epsilon$  or IRBP datasets considerably increases the NS scores above those of the most parsimonious trees. There are no putative synapomorphies in either  $\epsilon$  or IRBP sequences that would specifically group *Aotus* with *Callicebus*. The karyotypic variation in the cebids is considerable with *Aotus* varying from  $2n = 48$  to  $54$ , *Cebus* from  $2n = 52$  to  $54$  and *Saimiri* with a fairly stable karyotype of  $2n = 44$  (Assis and Barros, 1987; Saldanha, 1982).

#### Cebus-Saimiri clade

In our earlier analysis of  $\epsilon$ -globin gene sequences (Schneider et al., 1993), the consensus of most parsimonious trees grouped *Cebus*, *Saimiri*, *Aotus*, and the callitrichine clade as an unresolved tetrachotomy. However, after the  $\epsilon$  data set was enlarged with additional *Cebus* and *Saimiri* sequences (Harada et al., 1995), both  $\epsilon$  and IRBP sequences grouped these two genera (*Cebus* and *Saimiri*) together, corroborating traditional taxonomic classifications (see Rosenberger, 1981). Thus, even though the combined  $\epsilon$  and IRBP tree (Fig. 4) only modestly supports this *Cebus-Saimiri* group with a 94% bootstrap value and strength of grouping of 8, our provisional cladistic classification uses the subfamily name Cebinae for this presumed clade.

#### Callitrichine clade

The monophyly of the callitrichine clade (*Callithrix*, *Cebuella*, *Callimico*, *Leontopithecus*, *Saguinus*) with the inclusion of *Callimico* is no longer the subject of debate (Ford, 1986; Seuanez et al., 1988, 1989), although the phylogenetic relationships among genera are still controversial. Ford (1986) and Rosenberger (1981) consider *Leontopithecus* as the closest lineage to the *Callithrix-Cebuella* clade followed by *Saguinus* and then *Callimico* as the sister to the rest. On the other hand, Kay (1990) has

*Saguinus* as the sister of the *Callithrix-Cebuella* clade followed by *Leontopithecus* and then *Callimico*. The data obtained from  $\epsilon$  alone (Harada et al., 1995) or  $\epsilon$  and IRBP combined (Fig. 4), place *Callimico* as the sister of the *Cebuella-Callithrix* clade. It is important to emphasize here that while the node for the whole callitrichine clade is strongly supported by bootstrap and strength of grouping criteria the only node within the callitrichine clade that is strongly supported is the node for the *Callithrix-Cebuella* clade. *Callimico* appears as sister of the *Callithrix-Cebuella* clade when using  $\epsilon$ -globin gene sequences (Schneider et al., 1993; Harada et al., 1995), although without strong support from bootstrap and strength of grouping values. Even with the IRBP intron 1 sequences, *Callimico* appears related to the *Callithrix-Cebuella* clade, although first linked to *Leontopithecus* (Fig. 3). In both cases, the results contradict the morphological evidence such as reviewed by Kay (1994) that indicates the first ancestral separation within the callitrichines separates *Callimico* from all other callitrichines. Otherwise, our findings are in perfect agreement with those from Seuanez et al. (1989), obtained from Line-1 restriction data. Their results group *Callimico* with the *Callithrix-Cebuella* clade, next adds *Leontopithecus*, and finally *Saguinus*. These results also agree with the data of Sarich and Cronin (1980) on albumin immunological distances that place *Callimico* as sister of the *Callithrix-Cebuella* clade. Recently, Horovitz and Meyer (1995) sequenced a fragment of the 16S rRNA mitochondrial gene for most ceboid genera, and the most parsimonious tree that they obtained for these 16S sequences had *Callimico* as sister of the *Callithrix-Cebuella* clade. Seuanez et al. (1989), on comparing the karyotypes of *Callithrix jacchus*, *Callimico goeldii* and *Cebuella pygmaea*, concluded that *Callimico* is the closest lineage to the *Callithrix-Cebuella* clade.

#### Callithrix jacchus X Cebuella pygmaea

The divergence values obtained from IRBP nucleotide sequences for *Cebuella* and *Callithrix* are in the range of values for species of the same genus, corroborating our previous findings with the  $\epsilon$  gene (Schneider



et al., 1993), thus pointing to the phylogenetic closeness of *Cebuella pygmaea* to members of the genus *Callithrix*. Analyzing morphological data, both Moynihan (1976) and Rosenberger (1981) have suggested that *Cebuella* should be included in the genus *Callithrix*.

Recently, Nagamachi et al. (1992) observed that the karyotypes of *Cebuella pygmaea* ( $2n = 44$ ) and *Callithrix emiliae* ( $2n = 44$ ) differed from that of *Callithrix jacchus* ( $2n = 46$ ) by a pericentric inversion of chromosome 19 and a Robertsonian translocation (20/16 in *Cebuella* and 22/16 in *Callithrix emiliae*). *Cebuella* and the Amazonian *Callithrix emiliae* differ from each other by only a reciprocal translocation between an acrocentric autosome and the short arm of the submetacentric chromosome which distinguishes their karyotype from that of *Callithrix jacchus*. Seuanez et al. (1989) also observed a close karyotypic relationship between *Callithrix argentata* and *Cebuella*. In our opinion, there may now be enough evidence (morphological, cytogenetic and molecular) to place *Cebuella pygmaea* in the genus *Callithrix* (*Callithrix pygmaea*).

### Two or three families?

Our work with  $\epsilon$  sequences supports the view that the New World monkeys should be divided into two families: the Atelidae [containing subfamilies Atelinae (*Ateles*, *Brachyteles*, *Lagothrix* and *Alouatta*) and Pitheciinae (*Pithecia*, *Chiropotes*, *Cacajao* and *Callicebus*)], and the Cebidae (containing *Aotus*, *Cebus*, *Saimiri*, and the callitrichines). Conversely, the IRBP data alone place the pitheciines closer to the cebids, although without statistical support. IRBP and  $\epsilon$ , analyzed together, group pitheciines with atelines as suggested by the  $\epsilon$  data alone, but in this case with very low bootstrap and strength of grouping values. Although our results in balance do not rule out dividing the ceboids into the two families that we originally proposed (Schneider et al., 1993), they are also consistent with the three families proposed by Harada et al. (1995): Cebidae for cebines (*Cebus* and *Saimiri*), aotines (*Aotus*), and callitrichines; Pitheciidae for the pitheciines (*Callicebus* and pitheci-

ins); and Atelidae for the atelines (*Alouatta* and atelins).

### Concluding remarks

Despite many unanswered questions, our cladistic results from  $\epsilon$ -globin and intron 1 IRBP gene sequences provide important information on the phylogeny of New World monkeys. They definitely place *Callicebus* as a sister of pitheciins; they reinforce the monophyly of the cebids including the callitrichines, *Aotus*, *Cebus*, and *Saimiri*; they suggest a close relationship between *Cebus* and *Saimiri*; they support the reclassification of *Cebuella* as a member of the genus *Callithrix*, and finally they show a congruent branching pattern in the ateline clade, placing *Alouatta* as the oldest lineage and having *Brachyteles* and *Lagothrix* share the most recent ancestor. However, some points remain to be clarified, such as: (i) the exact branching pattern of *Cebus*-*Saimiri*, *Aotus*, and callitrichines, (ii) the precise positions of *Leontopithecus*, *Saguinus*, and *Callimico* within the callitrichines, and (iii) the phylogenetic position of pitheciines in relation to atelines and cebids. To clarify these points, it will be important to enlarge the DNA sequence data from additional ceboid species and outgroup species (from a broader range of noncebid primates) and from noncoding regions (introns) of additional nuclear genes.

### ACKNOWLEDGMENTS

We thank Dr. Eric Cabot for the sequence editor (ESEE200c), Dr. Jose Augusto Pereira Carneiro Muniz (Centro Nacional de Primatas-Belem, Para, Brazil), Dr. Ademar Coimbra-Filho and Dr. Alcides Pissinatti (Centro de Primatologia do Rio de Janeiro, RJ, Brazil), and Dr. Filomeno Encarnacion (Proyecto Peruano de Primatología, Iquitos, Peru) for the samples used in this paper. We also thank Drs. Steve Ferrara and Calvin Porter for assistance in revising the manuscript. This research was made possible thanks to grants from CNPq-Brazil (910043/91.4; 201142/91.0; 201596/92.0), FINEP-Brazil (6.6.94.0034.00), NSF-USA (DEB911-6098), and NIH-USA (HL33940). We also thank the Wayne State University Macro Molecular Core Facility for synthesis of the

primers for PCR amplification and sequencing.

### LITERATURE CITED

- Anthony MRL, and Kay FR (1993) Tooth form and diet in ateline and alouattine primates: Reflections on the comparative method. *Am. J. Sci.* 293:356–382.
- Assis MFL, and Barros R (1987) Karyotype pattern of *Saimiri ustus*. *Int. J. Primatol.* 8:432 (Abstract).
- Baba ML, Darga L, and Goodman M (1980) Biochemical evidence on the phylogeny of Anthropoidea. In R Ciochon and AB Chiarelli (eds.): *Evolutionary Biology of the New World Monkeys and Continental Drift*. New York: Plenum Press, pp 423–443.
- Barry D, and Hartigan JA (1987) Statistical analysis of hominoid molecular evolution. *Stat. Sci.* 2:191–210.
- Bremner K (1988) The limits of amino acid sequence data in angiosperm phylogenetic reconstruction. *Evolution* 42:795–803.
- Bunn HF, and Forget DG (1986) *Hemoglobin: Molecular, Genetics, and Clinical Aspects*. Philadelphia: Saunders.
- Cabot EL, and Beckenbach AT (1989) Simultaneous editing of multiple nucleic acid and protein sequences with ESEE. *Comp. Applic. Biosc.* 5:233–234.
- Cabrera A (1958) *Catálogo de los mamíferos de América del Sur*. *Rev. Mus. Argentino Cienc. Nat. Bernardino Rivadavia* 4:1–307.
- Cavalli-Sforza LL, and Edwards AWF (1967) Phylogenetic analysis: Models and estimation procedures. *Evolution* 21:550–570.
- Cronin JE, and Sarich VS (1975) Molecular systematics of the New World Monkeys. *J. Hum. Evol.* 4:357–375.
- Czelusniak J, and Goodman M (1995) Hominoid phylogeny estimated by model selection using goodness of fit significance tests. *Mol. Phylogenet. & Evol.* 4:283–290.
- De Boer LEM (1974) Cytotaxonomy of the Platyrrhini (Primates). *Genen. Phaenen.* 17:1–115.
- Dunlap SS, Thorington RW Jr., and Aziz MA (1985) Forelimb anatomy of New World monkeys: Myology, and the interpretation of primitive anthropoid models. *Am. J. Phys. Anthropol.* 68:499–517.
- Dutrillaux B (1988) Chromosome evolution in primates. *Folia Primatol* 50:134–135.
- Felsenstein J (1989) PHYLIP-Phylogeny inference package (version 3.2). *Cladistics* 5:164–166.
- Fleagle JG (1988) *Primate Adaptation & Evolution*. New York: Academic Press.
- Fong S-L, Fong W-B, Morris TA, Kedzie KM, and Bridges CD (1990) Characterization and comparative structural features of the gene for human interstitial retinol-binding protein. *J. Biol. Chem.* 265:3648–3653.
- Ford SM (1986) Systematics of the New World Monkeys. In DR Swindler and J Erwin, (eds.): *Comparative Primate Biology, Volume 1: Systematics, Evolution and Anatomy*. New York: Alan R. Liss, pp. 73–135.
- Goodman M, Olson CB, Beeber JE, and Czelusniak J (1982) New perspectives in the molecular biological analysis of mammalian phylogeny. *Acta. Zool. Fennica* 169:19–35.
- Goodman M, Czelusniak J, and Beeber JE (1985) Phylogeny of primates and other eutherian orders: A cladistic analysis using amino acid and nucleotide sequence data. *Cladistics* 1:171–185.
- Harada ML, Schneider H, Schneider MPC, Sampaio I, Czelusniak J, and Goodman M (1995) DAN evidence on the phylogenetic systematics of New World monkeys: support for the sister-grouping of *Cebus* and *Saimiri* from two unlinked nuclear genes. *Mol. Phylogenet. Evol.* 4:331–349.
- Hendy MD, and Penny D (1982) Branch and bound algorithms to determine minimal evolutionary trees. *Math. Biosc.* 59:277–290.
- Hershkovitz P (1977) *Living New World Monkeys*. Chicago: University of Chicago Press.
- Hershkovitz P (1990) Titis, New World Monkeys of the genus *Callicebus* (Cebidae, Platyrrhini): A preliminary taxonomic review. *Fieldiana Zool.* 55:1–109.
- Hill WCO (1962) *Primates, Comparative Anatomy and Taxonomy*. Vol. V. Cebidae, Part B. Edinburgh: University Press.
- Horowitz I, and Meyer A (1995) Systematics of New World monkeys (Platyrrhini, Primates) based on 16S mitochondrial DAN sequences: A comparative analysis of different weighting methods in cladistic analysis. *Mol. Phylogenet. Evol.* 4:448–456.
- Jukes TH and Cantor CR (1969) Evolution of protein molecules. In HN Munro (ed.): *Mammalian Protein Metabolism*. New York: Academic Press, pp. 21–32.
- Kay RF (1990) The phyletic relationships of extant and fossil Pitheciinae (Platyrrhini, Anthropoidea). *J. Human Evol.* 19:175–208.
- Kay RF (1994) “Giant” tamarin from the Miocene of Colombia. *Am. J. Phys. Anthropol.* 95:333–353.
- Martin RD (1990) *Primitive Origins and Evolution*. London: Chapman and Hill.
- Moyonhan M (1976) *The New World Primates: Adaptive Radiation and the Evolution of Social Behaviour, Languages, and Intelligence*. New Jersey: Princeton University Press.
- Nagamachi C, Pieczarka J, and Barros RM (1992) Karyotypic comparison among *Cebuella pygmaea*, *Callithrix jacchus* and *C. emiliae* (Callitrichidae, Primates) and its taxonomic implications. *Genetics* 85:249–257.
- Napier JR, and Napier PH (1967) *A Handbook of Living Primates*. New York: Academic Press.
- Navidi WC, Churchill GA, and von Haesler A (1991) Methods for inferring phylogenies from nucleic acid sequence data by using maximum likelihood and linear invariants. *Mol. Biol. Evol.* 8:128–143.
- Pieczarka J, Nagamachi CY, and Barros RMS (1989) The karyotype of *Ateles paniscus paniscus* (Cebidae, Primates): 2n = 32. *Rev. Brasil. Genet.* 12:543–551.
- Porter CA, Sampaio I, Schneider H, Schneider MPC, Czelusniak J, and Goodman M (1995) Evidence on primate phylogeny from  $\epsilon$ -globin gene sequences and flanking regions. *J. Mol. Evol.* 40:30–55.
- Powell MJD (1989a) The updating of matrices of conjugate directions in optimization algorithms. In DF Griffiths and GA Watson (eds.): *Numerical Analysis 1989*. Harlow, England: Longman Scientific and Technical, pp. 193–205.
- Powell MJD (1989b) TOLMIN: a Fortran package for linearly constrained optimization problems. Report DAMTP 1989/NA2, University of Cambridge.

- Rosenberger AL (1981) Systematics: The higher taxa. In AF Coimbra-Filho and R Mittermeier (eds.): *Ecology and Behaviour of Neotropical Primates*, Vol. 1. Rio de Janeiro: Academia Brasileira de Ciências, pp. 9–27.
- Rosenberger AL (1984) Fossil New World monkeys dispute the molecular clock. *J. Hum. Evol.* 13:737–742.
- Rosenberger AL, and Strier KB (1989) Adaptive radiation of the ateline primates. *J. Hum. Evol.* 18:717–750.
- Rosenberger AL, Setoguchi T, and Shigehara N (1990) The fossil record of callitrichine primates. *J. Hum. Evol.* 19:209–236.
- Saitou N, and Nei M (1987) The neighbor-joining method; a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* 4:406–425.
- Saldanha PH (1982) *Genética comparada de primatas Brasileiros*. São Paulo: Rev. Bras. Genet.
- Sambrook J, Fritsch EF, and Maniatis T (1989) *Molecular Cloning: A Laboratory Manual*, 2nd ed. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press.
- Sanger F, Nicklen S, and Coulson AR (1977) DNA sequencing with chain-termination inhibitors. *Proc. Natl. Acad. Sci. USA* 74:5463–5468.
- Sarich VM, and Cronin JE (1980) South American mammal molecular systematics, evolutionary clocks and continental drift. In R Ciochon and AB Chiarelli (eds.): *Evolutionary Biology of the New World Monkeys and Continental Drift*. New York: Plenum Press, pp. 399–421.
- Schneider H, Schneider MPC, Sampaio MIC, Harada ML, Stanhope M, and Goodman M (1993) Molecular phylogeny of the New World monkeys (Platyrrhini, Primates). *Mol. Phylogenet. Evol.* 2:225–242.
- Seuanez HN, Forman L, and Alves G (1988) Comparative chromosome morphology in three callitrichid genera: *Cebuella*, *Callithrix*, and *Leontopithecus*. *J. Hered.* 79:418–424.
- Seuanez HN, Forman L, Matayoshi T, and Fanning T (1989) The *Callimico goeldii* (Primates, Platyrrhini) genome: Karyology and middle repetitive (LINE-1) DNA sequences. *Chromosoma* 98:389–395.
- Simons EL (1972) *Primate Evolution: An Introduction to Man's Place in Nature*. New York: Macmillan.
- Simpson GG (1945) The principles of classification and classification of mammals. *Amer. Mus. Nat. Hist.* 85:1–350.
- Zingesser MR (1973) Dentition of *Brachyteles arachnoides* with reference to alouattine and ateline affinities. *Folia Primatol* 20:351–390.